

Geometrical optics without singularities: using the ray time as the coordinate space

I.Y. Dodin^{1,2,†}, N.A. Lopez³, Tingjing Xing^{2,4}, Rune Højlund Marholt⁵ and Valerian H. Hall-Chen^{4,6}

¹Princeton Plasma Physics Laboratory, Princeton, NJ 08543, USA

²Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

³Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford OX1 3PU, UK

⁴Institute of High Performance Computing, Agency for Science, Technology and Research (A*STAR), Singapore 138632, Republic of Singapore

⁵Section for Plasma Physics and Fusion Energy, Department of Physics, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

⁶Future Energy Acceleration and Translation (FEAT), Strategic Research and Translational Thrust (SRTT), A*STAR, Singapore 138632, Republic of Singapore

Corresponding author: I.Y. Dodin, iododin@sjtu.edu.cn

(Received 9 September 2025; revision received 30 December 2025; accepted 11 January 2026)

Geometrical optics (GO) is widely used for reduced modelling of waves in plasmas, but it fails near reflection points, where it predicts a spurious singularity of the wave amplitude. We show how to avoid this singularity by adopting a different representation of the wave equation. Instead of the physical coordinate x and the wavevector k , we use the ray time τ as the new canonical coordinate and the ray energy h as the associated canonical momentum. To derive the envelope equation in the τ -representation, we construct the Weyl symbol calculus on the (τ, h) space and show that the corresponding Weyl symbols are related to their (x, k) counterparts by the Airy transform. This allows us to express the coefficients in the envelope equation through the known properties of the original dispersion operator. When necessary, solutions of this equation can be mapped to the x -space using a generalised metaplectic transform. However, the field *per se* might not even be needed in practice. Instead, knowing the corresponding Wigner function usually suffices for linear and quasilinear calculations. As a Weyl symbol itself, the Wigner function can be mapped analytically, using the aforementioned Airy transform. We show that the standard Airy patterns that form in regions where conventional GO fails are successfully reproduced within metaplectic GO (MGO) simply by remapping the field from the τ -space to the x -space. An extension to mode-converting waves is also presented. This formulation, which we call generalised MGO, can be particularly useful, for example, for reduced modelling of the O–X conversion in inhomogeneous plasma near the critical density, an effect that is important for fusion applications and also occurs in the ionosphere. Overall, MGO can replace GO for any practical purposes, because it better handles cutoffs and is similar otherwise.

Key words: plasma waves

[†]Present address: Shanghai Jiao Tong University, Shanghai, China.

1. Introduction

1.1. Background

The geometrical-optics (GO) approach is widely used for reduced modelling of waves in plasmas and other media (Kravtsov & Orlov 1990; Tracy *et al.* 2014). In this approach, instead of solving the wave equations *per se*, one converts them into Hamilton's equations for wave rays and a first-order (in time or the ray path) differential equation for the wave envelope. Then, one calculates the ray trajectories that are relevant to the initial conditions ('reference rays') and solves the envelope equations on those trajectories. As an initial-value algorithm, this scheme is computationally inexpensive and has also yielded a number of spinoffs, such as beam tracing (Pereverzev 1998; Poli *et al.* 2018; Hall-Chen, Parra & Hillesheim 2022) and quasioptics, which can account for both transverse diffraction (Balakin *et al.* 2007) and mode conversion (Dodin *et al.* 2019; Yanagihara *et al.* 2021).¹

Still, the GO approach is limited in that it assumes that the characteristic scale L of the medium is much larger than the local wavelength λ . Even for high-frequency waves, such as those in the electron-cyclotron range, this assumption unavoidably breaks down near cutoffs (reflection points), where λ/L is infinite and the conventional GO predicts a spurious singularity of the field amplitude. For example, this is an issue in modelling the O–X mode conversion near the critical plasma density, which is a part of a promising scheme for heating overdense plasmas in fusion reactors (Preinhaelter & Kopecký 1973; Hansen, Lynov & Michelsen 1985) and also occurs in the ionosphere (Mjølhus 1990; Eliasson & Papadopoulos 2016). Since those waves experience many oscillations during the mode-conversion process, one can expect that some GO-like model should be possible for them, even though the conversion occurs near a cutoff. This motivates a search for alternative formulations of GO that do not rely on the smallness of λ/L *per se*.

One such formulation, which has long been known, is based on Maslov's approach (Maslov & Fedoriuk 1981; Ziolkowski & Deschamps 1984; Thomson & Chapman 1985). The idea of this approach is based on the fact that a wave experiencing reflection in the physical space satisfies the conditions of GO in the spectral space, so one can reinstate GO locally by applying the (spatial) Fourier transform. However, applying Maslov's method is cumbersome in that the exact moment for performing the Fourier transform is only loosely specified and thus codes have to be supervised. The more recent methods (Littlejohn 1985, 1986; Kay 1994a; Zor & Kay 1996; Alonso & Forbes 1997, 1999; Madhusoodanan & Kay 1998) remedy this issue by using continuous transformations of the wavefield along ray trajectories. However, these methods are restricted in the types of solutions sought and are usually applicable only to specific equations, such as the Schrödinger equation for quantum particles. Practical modelling of waves in plasmas, especially magnetised plasmas, requires a more general framework.

The first framework reasonably fit for plasma applications, called metaplectic geometrical optics (MGO), was developed by Lopez & Dodin (2019, 2020, 2021a,b) and Donnelly, Lopez & Dodin (2021); for an overview, see Lopez & Dodin (2022). Its basic idea is as follows: instead of propagating the wavefield in the physical space, one propagates it, by integrating GO equations over a differentially small time, in a space that is locally tangent to the ray trajectory in the ray phase space. Then, one remaps the field to the space tangent to the reference ray at the next location

¹In its standard formulation, GO is also closely related to the Wentzel–Kramers–Brillouin (WKB) approximation, so we will use the terms 'conventional GO' and 'WKB' interchangeably.

on the ray and iterates. The remapping is done using metaplectic transforms (MTs), a subject that we will introduce shortly (§ 1.2). An advantage of this formulation is that it can be used for any linear wave equation, including integro-differential equations that emerge in kinetic wave theory. It also readily incorporates transverse diffraction and mode conversion and has been proven workable in calculations of the field caustics within various wave models (Lopez & Dodin 2021*b*). That said, the need for constant remapping of the wavefield between neighbouring tangent spaces remains an inconvenience in MGO. It would be better to have a differential equation that would describe the continuous evolution of the wave amplitude on reference rays themselves, while mapping the resulting solutions to the physical space would remain an optional diagnostic. The purpose of this paper is to propose such a formulation of MGO. (In the following, the term ‘MGO’ refers to the new formulation that we report here.)

Specifically, the reformulation of MGO reported here is as follows. We adopt the ray time τ as the new coordinate and the ray energy h as the new canonical momentum (wavenumber). Unlike any given tangent space, which allows reinstating GO locally, the variables (τ, h) are convenient globally. Reflection never occurs in the τ -space; therefore, the GO wave field remains non-singular in the τ -representation indefinitely. We derive a differential equation for this field in the τ -space using the symplectic curvature of the phase-space trajectory as a small parameter. We also propose various tools for remapping the wave field, and quadratic functionals thereof, to the physical space x . This remapping is needed only for initialising the field in the τ -space and for outputting its final state to the x -space. Propagation, though, can be done entirely in the τ -space, at the same low computational cost as that of the traditional GO in its validity domain.

But what does it mean to propagate field in the τ -space? Answering this question requires introducing some mathematical concepts, most of which are not new but are not commonly used in plasma theory. We do this briefly in § 1.2 (more details, including some new theorems, are presented in later sections) and return to our main results in § 1.3. An outline of the paper is presented in § 1.4.

1.2. A brief overview of the mathematical tools used in MGO

Suppose that one has some field $\psi_{\hat{x}}$ on a coordinate space x . (The reason for adding the index \hat{x} will become clear shortly.) Any function $\hat{U}\psi_{\hat{x}}$, where \hat{U} is a unitary transformation, contains just as much information as $\psi_{\hat{x}}$ itself. Then, it makes sense to consider the set of all possible $\hat{U}\psi_{\hat{x}}$ as a single object, called ket vector $|\psi\rangle$, and each $\hat{U}\psi_{\hat{x}}$ as a representation of $|\psi\rangle$. For example, $\psi_{\hat{x}}(x)$ itself can be understood as a projection of $|\psi\rangle$ on the eigenvector of the position operator \hat{x} that corresponds to the eigenvalue x ; i.e. $\psi_{\hat{x}}(x) = \langle e_{\hat{x}}(x) | \psi \rangle$. Performing a ‘coordinate transformation’, i.e. choosing some \hat{q} as the new position operator instead of \hat{x} , induces a q -representation $\psi_{\hat{q}}(q) = \langle e_{\hat{q}}(q) | \psi \rangle$, where $e_{\hat{q}}(q)$ is the eigenvector of \hat{q} corresponding to the eigenvalue q . It is connected to the original x -representation via a unitary integral operator, $\psi_{\hat{q}}(q) = \int M(q, x) \psi_{\hat{x}}(x) dx$, with kernel $M(q, x) = \langle e_{\hat{q}}(q) | e_{\hat{x}}(x) \rangle$. The mapping between $\psi_{\hat{x}}$ and $\psi_{\hat{q}}$ is an MT. In this paper, we introduce more than two position operators and, thus, multiple MTs as well.

One example of an MT is the Fourier transform, which maps functions in the x -representation to those in the wavevector, or k -representation, i.e. corresponds to $\hat{q} = \hat{k}$. (Because the x -representation of the wavevector operator is $\hat{k}_{\hat{x}} = -i\partial_x$, this

corresponds to $M(q, x) \propto e^{-iqx}$.) MTs are also used to introduce mixed representations, where $\hat{q} = A\hat{x} + B\hat{k}$, with constant A and B .² In this sense, MTs are a natural generalisation of the Fourier transform and are well studied as such (Littlejohn 1986). However, linear operator transformations (also used in the earlier formulation of MGO) are not enough for our purposes here. Since we aim to solve for the wavefield in the coordinate frame aligned with the ray, it will be necessary to consider the possibility of nonlinear dependence of \hat{q} on \hat{x} and \hat{k} . The corresponding MTs have been discussed in the quantum-physics literature to some extent (for example, see Mello & Moshinsky 1975), and their semiclassical limit has attracted interest in the area of atomic and molecular dynamics (Miller 1974). However, said literature is concerned with problems different from those in which are of interest here, so we basically start from scratch in this paper and will mention parallels with the existing literature only where they are relevant.

MTs are closely related to the Weyl symbol calculus and Wigner functions (§ 2). For any given operator \hat{A} expressed through \hat{x} and \hat{k} , its (Weyl) symbol A_z is a function of $(x, k) \equiv \mathbf{z}$ that is typically close, in regimes in which we are interested, to the one that is obtained from \hat{A} by replacing each instance of \hat{x} and \hat{k} with x and k , respectively.³ (In the coordinate representation, this means replacing just ∂_x with ik .) In particular, the symbol of $|\psi\rangle\langle\psi|$ ⁴ on the space \mathbf{z} is called the Wigner function of $\psi_{\hat{x}}$ and equals the Fourier image of $\psi_{\hat{x}}(x + s/2)\psi_{\hat{x}}^*(x - s/2)$ with respect to s , assuming $\psi_{\hat{x}}$ is a scalar.

One can also use base operators other than \hat{x} and \hat{k} , say, $\hat{\tau}$ and \hat{h} . Then, symbols, including Wigner functions, become functions of $(\tau, h) \equiv \mathbf{r}$, respectively. If the relation between (\hat{x}, \hat{k}) and $(\hat{\tau}, \hat{h})$ is linear, the new symbols are related to the old ones via a simple formula $A_r(\mathbf{r}) = A_z(\mathbf{z}(\mathbf{r}))$, i.e. exhibit symplectic invariance; however, this is generally not the case for nonlinear transformations. A firm grasp on these facts is important to understanding the following section.

1.3. Main results

Developing the new formulation of MGO as described in § 1.1 means answering the following questions. (i) What is the natural position operator $\hat{\tau}$ for MGO? (ii) How does one write GO equations in the τ -representation? (iii) What is the small parameter that replaces the GO parameter λ/L ? (iv) How does one map a solution from the τ -representation to the x -representation? Here, we summarise how we answer these questions in the rest of the paper.

Suppose a wave is governed by $\hat{H}|\psi\rangle = 0$, where \hat{H} is a Hermitian operator. (A small anti-Hermitian addition to \hat{H} can be easily accommodated as a perturbation.) A systematic derivation of the conventional GO from this equation (§ 4) starts with introducing the Weyl symbol of \hat{H} , denoted H_z , where the index \mathbf{z} indicates that the corresponding Weyl calculus is constructed on the original ray phase-space $(x, k) \equiv \mathbf{z}$. Then, one Taylor-expands H_z around the ray with λ/L as a small parameter and converts this approximate symbol into a new operator, which

²In the literature, the term ‘MT’ is typically reserved just for this particular case, where \hat{q} is linear in both \hat{x} and \hat{k} . At the risk of irritating our readers, we will ignore this restrictive definition and use the term ‘MT’ for the transformation of the field representation for any $\hat{q}(\hat{x}, \hat{k})$.

³For example, the susceptibility $\chi(x, k)$ of an inhomogeneous medium is, at least approximately, the Weyl symbol of the response operator $\hat{\chi}$ that enters Maxwell’s equations. (Depending on a problem, x can be a timespace coordinate; then k also includes the frequency.)

⁴Up to a constant coefficient that depends on a convention.

serves as an approximation of \hat{H} . This leads to an approximate amplitude equation on a reference ray. Assuming that H_z is a scalar (vector waves are discussed later), the reference ray itself is governed by Hamilton's equations, with H_z serving as the Hamiltonian.

In MGO, we introduce a different phase space, $(\tau, h) \equiv \mathfrak{r}$, which is aligned with the ray (§ 5). It turns out that the natural position operator $\hat{\tau}$ in this case is the one that is associated with the momentum operator $\hat{h} = \hat{H}$ via the canonical commutation relation $[\hat{\tau}, \hat{h}] = i$,⁵ i.e. τ is the ray time. We introduce the Weyl calculus on the space \mathfrak{r} and consider the corresponding symbol $H_{\mathfrak{r}}$. MGO envelope equations are derived much like in GO, via Taylor-expanding $H_{\mathfrak{r}}$ (§ 6). However, the relevant small parameter in this case is not λ/L but $\epsilon \doteq (\Delta x \Delta k)^{-1}$, where Δx and Δk are the characteristic scales of the original symbol H_z along the coordinate axis and the wavenumber axis, correspondingly.⁶ This parameter, which was not articulated in the previous version of MGO for the lack of a differential formulation, now explicitly shows up in the envelope equations (see later). It can be understood as the squared symplectic curvature of the ray trajectory, or, at least, as a characteristic value thereof. Note that, unlike for GO, neither the absolute value of k nor the group velocity in the x -space are important for MGO.

How does one find $H_{\mathfrak{r}}$? In practice, one generally knows \hat{H} in the x -representation. This means that one knows H_z . To obtain MGO equations explicitly, one needs to express $H_{\mathfrak{r}}$ through H_z . Since the mapping between \mathfrak{z} and \mathfrak{r} is nonlinear, exact symplectic invariance is not to be expected, so we derive the corresponding relation from scratch. We show that the mapping between the symbols of a given operator in any two representations is determined by the Wigner function of the MT kernel M . We also show that, for the $\mathfrak{z} \mapsto \mathfrak{r}$ transformation at small ϵ in particular, M is delta-shaped and this mapping can be approximated with the Airy transform. This finding is critical for MGO in multiple aspects. In particular, for smooth symbols like H_z , it leads to $H_{\mathfrak{r}}(\mathfrak{r}) = H_z(\mathfrak{z}(\mathfrak{r})) + \mathcal{O}(\epsilon^2)$ even for our delta-shaped M . This is an important new result in that, unlike $\mathcal{O}(\epsilon)$, corrections $\mathcal{O}(\epsilon^2)$ are negligible within the accuracy needed for MGO. In other words, in MGO, one can ignore the difference between $H_{\mathfrak{r}}(\mathfrak{r})$ and $H_z(\mathfrak{z}(\mathfrak{r}))$ notwithstanding the lack of exact symplectic invariance. This makes the MGO envelope equation as simple as the GO envelope equation. The only difference between the two is that the MGO equation is formulated in a new space τ and H_z is differentiated with respect to different variables defined such that singularities are avoided (§ 6).

We also propose new ways to approximate the MT kernel itself, which then enables mapping solutions of the MGO envelope equation to the physical space x . Related problems have been traditionally attracting attention in the quantum mechanics literature. However, they are not particularly important in the context of modelling waves in plasmas for several reasons. First of all, possible errors caused by inaccuracy of the mapping $\psi_{\hat{q}} \mapsto \psi_{\hat{x}}$ do not accumulate along the ray trajectory, so there is no need to make this mapping particularly precise. The wavefield is propagated in MGO in a single space, τ , and mapping to the x -space is just an optional diagnostic. Second, although wave simulations are traditionally expected to yield spectacular illustrations, the field in the x -space does not actually matter much for practical

⁵Notably, this τ is also basically what 't Hooft (2024) calls the ontological variable in a quantum-mechanical context.

⁶If $\hat{\tau}$ coincides with \hat{x} , when MGO is identical to GO, then $\Delta x \sim L$ and $\Delta k \sim k \sim 1/\lambda$, so ϵ becomes the usual GO parameter λ/L .

purposes. One might not even need to leave the τ -space except for initialising the field and also calculating the final state, and those mappings may be easy to apply. Third, even when spatial profiles are actually needed, they are not of the field *per se*, but rather of quadratic functionals thereof. Such functionals can always be expressed through the Wigner function of the wavefield. (In fact, any reasonably general quasilinear calculations are expressed more naturally through Wigner functions than through fields *per se*; see Dodin (2022, 2024).) As a Weyl symbol, a Wigner function can be remapped from the r -space to the z -space analytically, using the aforementioned Airy transform or its asymptotics. We show that this allows one to easily recover the well-known WKB results, as well as the also well-known Airy profiles of wavefields near cutoffs (§ 6), which are usually considered as an epitome of GO failure. In a way, the Airy function is a cornerstone of MGO, because, basically, any wavefield locally looks like an Airy field in a certain representation.

We then show that these results largely extend to vector waves as well (§ 7). Since H_z is a matrix in this case, reference rays and \hat{h} are constructed out of its relevant (small) eigenvalues rather than H_z *per se*. We also derive equations that capture mode conversion when H_z is degenerate and illustrate their application on a simple example. Finally, we introduce the new concept of a metaplectic resonance, which generalises the concept of the Cherenkov resonance to waves that are quasimonochromatic in their natural τ -representation, but not necessarily in the physical x -representation (§ 8). Applications of this theory, as well as the (potentially straightforward) extension of MGO to quasioptical beams are left to future work.

In summary, the main results of this paper are: the asymptotic formulae for the Weyl-symbol transformations (most importantly, (5.65) and (5.66)), the envelope equations for scalar and vector waves, the idea of metaplectic remapping of Wigner functions of wave fields instead of the wave fields themselves, and the formulation of the concept of metaplectic resonance.

1.4. Outline

The rest of this paper is organised as follows. In § 2, we introduce the Weyl symbol calculus and Wigner functions. This material is not new and is included mostly to introduce some notation. This is largely where the tutorial ends, although we do revisit known results throughout the paper to put our formulation in the context of common knowledge. We do so extensively because this paper is intended to form a foundation for new MGO studies based on nonlinear phase-space transformations⁷ and it is important to leave no stone unturned at this stage. It is also our goal here to present the logic and the thought process beyond the main results *per se*, as we hope that, in the future, MGO will be not only applied, but also developed further.

In § 3, we formalise the concept of an MT, derive the key properties of MTs and present examples. Some parts of this material can be found in the literature. However, MTs corresponding to nonlinear phase-space transformations are not covered in the literature to an extent sufficient, or in a context suitable, for developing MGO. Section 3 closes the gap and, in this sense, is new. In § 4, we restate the conventional GO, which is used as an important reference point in the rest of the paper. The final equations are not new; however, new is the finding that the well-known GO envelope equation can be understood as a result of applying a particular MT to a full-wave equation. This readily suggests that other representations of the

⁷Our earlier papers on MGO used only linear phase-space transformations and were not concerned with what we focus on later. Those papers are neither built upon nor reviewed here.

envelope equation are also possible. We also show how the MT kernel M itself can be approximated using GO methods in some cases and reproduce some known results as a special case. In §5, we derive $\hat{\tau}$, \hat{h} , M and its Wigner function μ for mapping symbols between the z -space and the r -space. Section 5.4.3 summarises our main new findings regarding symbol transformations. We use these results to formulate MGO, but they can also have broader applications.

Our new results specific to MGO are presented in §§6–8. In §6, we derive the MGO envelope equation for scalar waves and discuss how to map its solutions to the x -space. In §7, we generalise MGO to vector waves and discuss mode conversion in particular. Finally, in §8, we introduce the concept of a metaplectic resonance.

Auxiliary calculations are presented in the appendices. In Appendix A, we rederive the ray equations, both for completeness and also to introduce some terminology. In Appendix B, we discuss the possibility of introducing the local angle–action variables, and the associated operators, for unbounded GO trajectories. In Appendix C, we discuss some properties of the auxiliary functions associated with the aforementioned Airy transform. In Appendix D, we rederive the Einstein–Brillouin–Keller quantisation within MGO, where it becomes particularly natural. In Appendix E, we derive an expression for the electromagnetic dissipation power per phase-space volume (as opposed to the spatial volume, as usual). Finally, in Appendix F, we summarise our main notation.

2. A maths primer

2.1. Notation

For any given Hermitian operator \hat{A} on a given Hilbert space \mathbb{H} , one can introduce a complete set of its orthogonal eigenvectors $|\mathbf{e}_{\hat{A}}(\lambda)\rangle \in \mathbb{H}$,

$$\hat{A} |\mathbf{e}_{\hat{A}}(\lambda)\rangle = \lambda |\mathbf{e}_{\hat{A}}(\lambda)\rangle, \quad \lambda = \lambda^*, \quad (2.1)$$

where λ are the corresponding eigenvalues. (We use $*$ to denote complex conjugation, so the second equation states that the eigenvalues of a Hermitian operator are real.) Assuming that the eigenspace of \hat{A} is $\mathbb{R} \equiv (-\infty, \infty)$,⁸ its eigenvectors can always be chosen such that

$$\langle \mathbf{e}_{\hat{A}}(\lambda') | \mathbf{e}_{\hat{A}}(\lambda'') \rangle = \delta(\lambda' - \lambda''), \quad (2.2)$$

where δ is the Dirac delta function. Then, they satisfy the identity

$$\int d\lambda |\mathbf{e}_{\hat{A}}(\lambda)\rangle \langle \mathbf{e}_{\hat{A}}(\lambda)| = \hat{1}, \quad (2.3)$$

where $\hat{1}$ is a unit operator. (Integrals in this paper are taken over \mathbb{R} except where specified otherwise.) Accordingly, \hat{A} can be expressed as

$$\begin{aligned} \hat{A} &= \int d\lambda' |\mathbf{e}_{\hat{A}}(\lambda')\rangle \langle \mathbf{e}_{\hat{A}}(\lambda')| \hat{A} \int d\lambda'' |\mathbf{e}_{\hat{A}}(\lambda'')\rangle \langle \mathbf{e}_{\hat{A}}(\lambda'')| \\ &= \int d\lambda' d\lambda'' |\mathbf{e}_{\hat{A}}(\lambda')\rangle \lambda'' \langle \mathbf{e}_{\hat{A}}(\lambda') | \mathbf{e}_{\hat{A}}(\lambda'') \rangle \langle \mathbf{e}_{\hat{A}}(\lambda'')| \\ &= \int d\lambda |\mathbf{e}_{\hat{A}}(\lambda)\rangle \lambda \langle \mathbf{e}_{\hat{A}}(\lambda)|. \end{aligned} \quad (2.4)$$

⁸A generalisation to \mathbb{R}^n is straightforward (Dodin 2022). Also, strictly speaking, one should introduce a rigged Hilbert space to allow for (2.2), but we do not strive for ultimate rigour here.

One can also use the basis of $|\mathbf{e}_{\hat{A}}\rangle$ to define the trace of any given operator \hat{O} :

$$\text{tr } \hat{O} \doteq \int d\lambda \langle \mathbf{e}_{\hat{A}}(\lambda) | \hat{O} | \mathbf{e}_{\hat{A}}(\lambda) \rangle. \quad (2.5)$$

Since the trace is invariant under the change of basis, any \hat{A} generates the same $\text{tr } \hat{O}$.

Among such operators on \mathbb{H} , let us assign some operator \hat{x} as the position operator. For any given vector $|\psi\rangle$ in \mathbb{H} , the eigenvectors of \hat{x} define the x -representation of $|\psi\rangle$, or a field $\psi_{\hat{x}}(x)$, via

$$\psi_{\hat{x}}(x) \doteq \langle \mathbf{e}_{\hat{x}}(x) | \psi \rangle. \quad (2.6)$$

Correspondingly, for any given operator \hat{O} , the field associated with $\hat{O} |\psi\rangle$ is as follows:

$$(\hat{O}_{\hat{x}} \psi_{\hat{x}})(x) = \langle \mathbf{e}_{\hat{x}}(x) | \hat{O} | \psi \rangle. \quad (2.7)$$

Using (2.3) for $\hat{A} = \hat{x}$, one can rewrite this as

$$\begin{aligned} (\hat{O}_{\hat{x}} \psi_{\hat{x}})(x) &= \int dx' \langle \mathbf{e}_{\hat{x}}(x) | \hat{O} | \mathbf{e}_{\hat{x}}(x') \rangle \langle \mathbf{e}_{\hat{x}}(x') | \psi \rangle \\ &= \int dx' O_{\hat{x}}(x, x') \psi_{\hat{x}}(x'), \end{aligned} \quad (2.8)$$

where the kernel $O_{\hat{x}}(x, x') \doteq \langle \mathbf{e}_{\hat{x}}(x) | \hat{O} | \mathbf{e}_{\hat{x}}(x') \rangle$ is the ‘matrix element’ of \hat{O} in the x -representation, or simply the x -representation of \hat{O} . In the special case when $\hat{O}_{\hat{x}}$ is a local or differential operator (that is, if $(\hat{O}_{\hat{x}} \psi_{\hat{x}})(x)$ can be expressed through $\psi_{\hat{x}}$ and its derivatives $\psi_{\hat{x}}^{(m)}$ with finite m), then $O_{\hat{x}}(x, x')$ can be expressed through the delta function, $\delta(x - x')$, and its corresponding derivatives, $\delta^{(m)}(x - x')$.

Using (2.4), one can further use this function to express \hat{O} as follows:

$$\hat{O} = \int dx' dx'' |\mathbf{e}_{\hat{x}}(x')\rangle O_{\hat{x}}(x', x'') \langle \mathbf{e}_{\hat{x}}(x'')|, \quad (2.9)$$

and its Hermitian adjoint as

$$\hat{O}^\dagger = \int dx' dx'' |\mathbf{e}_{\hat{x}}(x'')\rangle O_{\hat{x}}^*(x'', x') \langle \mathbf{e}_{\hat{x}}(x')|. \quad (2.10)$$

By definition,

$$x_{\hat{x}}(x, x') = \langle \mathbf{e}_{\hat{x}}(x) | \hat{x} | \mathbf{e}_{\hat{x}}(x') \rangle = x' \langle \mathbf{e}_{\hat{x}}(x) | \mathbf{e}_{\hat{x}}(x') \rangle = x' \delta(x - x') = x \delta(x - x'). \quad (2.11)$$

Then, $(\hat{x}_{\hat{x}} \psi_{\hat{x}})(x) = x \psi_{\hat{x}}(x)$, so

$$\hat{x}_{\hat{x}} = x. \quad (2.12)$$

Let us also introduce the wavevector operator \hat{k} as a generator of spatial translations:

$$\exp(-ia\hat{k}) |\mathbf{e}_{\hat{x}}(x)\rangle = |\mathbf{e}_{\hat{x}}(x + a)\rangle \quad (2.13)$$

(here, a is any real constant), whence

$$\begin{aligned}
 (\exp(-ia\hat{k}_{\hat{x}})\psi_{\hat{x}})(x) &= \langle \mathbf{e}_{\hat{x}}(x) | \exp(-ia\hat{k}) | \psi \rangle \\
 &= \langle \psi | \exp(ia\hat{k}) | \mathbf{e}_{\hat{x}}(x) \rangle^* \\
 &= \langle \psi | \mathbf{e}_{\hat{x}}(x-a) \rangle^* \\
 &= \psi_{\hat{x}}(x-a).
 \end{aligned} \tag{2.14}$$

At $a \rightarrow 0$, one has $\exp(-ia\hat{k}_{\hat{x}}) \rightarrow 1 - ia\hat{k}_{\hat{x}}$ and $\psi_{\hat{x}}(x-a) \rightarrow \psi_{\hat{x}}(x) - a \partial_x \psi_{\hat{x}}(x)$, so (2.14) leads to $-i\hat{k}_{\hat{x}}\psi_{\hat{x}}(x) \rightarrow -\partial_x \psi_{\hat{x}}(x)$ for any $\psi_{\hat{x}}(x)$. Thus,

$$\hat{k}_{\hat{x}} = -i\partial_x \tag{2.15}$$

and $k_{\hat{x}}(x, x') = -i\delta'(x - x')$. (The prime in δ' denotes the derivative with respect to the argument.) Such \hat{x} and \hat{k} are called Schrödinger operators, and here we will also refer to them as base operators. Note that the base operators satisfy the canonical commutation relation

$$[\hat{x}, \hat{k}] = i. \tag{2.16}$$

The converse is also true: if given Hermitian operators \hat{x} and \hat{k} satisfy (2.16), then there exists a representation, called the x -representation, in which (2.12) and (2.15) are satisfied; see, e.g. Galindo & Pascual (1990, § 2.12).

For $\xi_k(x) \doteq \langle \mathbf{e}_{\hat{x}}(x) | \mathbf{e}_{\hat{k}}(k) \rangle$, which is the x -representation of an eigenvector of \hat{k} , (2.15) yields $-i\partial_x \xi_k = k\xi_k$. Assuming the same normalisation as in (2.2), one obtains⁹

$$\langle \mathbf{e}_{\hat{x}}(x) | \mathbf{e}_{\hat{k}}(k) \rangle = \frac{1}{\sqrt{2\pi}} e^{ikx}, \tag{2.17}$$

where we chose the integration constant such that $\xi_k(0)$ is real. This convention will also be assumed from now on for the eigenfunctions of all other base operators.

Notably, (2.17) means that the k -representation of any given $|\psi\rangle$ is the Fourier transform of the corresponding $\psi_{\hat{x}}$:

$$\psi_{\hat{k}} \doteq \langle \mathbf{e}_{\hat{k}}(k) | \psi \rangle = \int dx \langle \mathbf{e}_{\hat{k}}(k) | \mathbf{e}_{\hat{x}}(x) \rangle \langle \mathbf{e}_{\hat{x}}(x) | \psi \rangle = \frac{1}{\sqrt{2\pi}} \int dx e^{-ikx} \psi_{\hat{x}}(x). \tag{2.18}$$

Obviously, the k -representation of \hat{k} is $\hat{k}_{\hat{k}} = k$. The k -representation of \hat{x} is found from

$$\begin{aligned}
 (\hat{x}_{\hat{k}}\psi_{\hat{k}})(k) &= \int dk' \langle \mathbf{e}_{\hat{k}}(k) | \hat{x} | \mathbf{e}_{\hat{k}}(k') \rangle \langle \mathbf{e}_{\hat{k}}(k') | \psi \rangle \\
 &= \int dx' dk' \langle \mathbf{e}_{\hat{k}}(k) | \hat{x} | \mathbf{e}_{\hat{x}}(x') \rangle \langle \mathbf{e}_{\hat{x}}(x') | \mathbf{e}_{\hat{k}}(k') \rangle \langle \mathbf{e}_{\hat{k}}(k') | \psi \rangle \\
 &= \int dx' dk' x' \langle \mathbf{e}_{\hat{x}}(x') | \mathbf{e}_{\hat{k}}(k) \rangle^* \langle \mathbf{e}_{\hat{x}}(x') | \mathbf{e}_{\hat{k}}(k') \rangle \langle \mathbf{e}_{\hat{k}}(k') | \psi \rangle \\
 &= \int dk' \langle \mathbf{e}_{\hat{k}}(k') | \psi \rangle \frac{1}{2\pi} \int dx' x' e^{i(k'-k)x'} \\
 &= i\partial_k \int dk' \langle \mathbf{e}_{\hat{k}}(k') | \psi \rangle \delta(k' - k) \\
 &= i\partial_k \psi_{\hat{k}}(k).
 \end{aligned} \tag{2.19}$$

⁹For n -dimensional x -space, here and further, one should replace 2π with $(2\pi)^n$.

Then, $\hat{x}_{\hat{k}} = i\partial_k$, so $-\hat{x}$ can be understood as a generator of translations in the k -space.

We will also use the notation

$$\hat{z} \equiv \begin{pmatrix} \hat{x} \\ \hat{k} \end{pmatrix} \equiv (\hat{x}, \hat{k})^\top, \quad \mathbf{z} \equiv \begin{pmatrix} x \\ k \end{pmatrix} \equiv (x, k)^\top, \quad (2.20)$$

where $^\top$ denotes transposition. In other words, $\mathbf{z} \equiv (x, k)^\top$ will be understood as a column vector and $\mathbf{z}^\top \equiv (x, k)$ will be generally understood as a row vector. (We ignored the difference between the two in the introduction for simplicity.) The upper and lower indices are not distinguished in this paper, but the Einstein summation convention for repeated indices is assumed. For two vectors $\mathbf{z}_1 = (a_1, b_1)^\top$ and $\mathbf{z}_2 = (a_2, b_2)^\top$, one has $\mathbf{z}_1^\top \mathbf{z}_2 = a_1 a_2 + b_1 b_2 \equiv \mathbf{z}_1 \cdot \mathbf{z}_2 = \mathbf{z}_2 \cdot \mathbf{z}_1$.¹⁰ We will also need a wedge product defined as

$$\mathbf{a} \wedge \mathbf{b} \doteq \mathbf{a} \cdot \mathbf{J} \mathbf{b} = -\mathbf{b} \wedge \mathbf{a}, \quad (2.21)$$

and \mathbf{J} is the canonical symplectic form, which satisfies $\mathbf{J}\mathbf{J} = -\mathbf{I}$:

$$\mathbf{J} \doteq \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathbf{I} \doteq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (2.22)$$

2.2. Weyl symbol calculus

Any given operator \hat{O} can be expressed as a linear superposition of non-commuting phase-space translations $\hat{T}_{\hat{\mathbf{a}}} \doteq e^{-i\hat{\mathbf{a}} \wedge \hat{\mathbf{z}}}$. Specifically, this can be expressed as follows:¹¹

$$\hat{O} = \frac{1}{2\pi} \int d^2\mathbf{z} O(\mathbf{z}) \hat{\Delta}_{\mathbf{z}}, \quad O(\mathbf{z}) = \text{tr}(\Delta_{\mathbf{z}} \hat{A}), \quad (2.23)$$

where the weight function $O(\mathbf{z}) \equiv \text{symb } \hat{O}$ is called the Weyl symbol (or just ‘symbol’¹²) of \hat{O} and $\hat{\Delta}_{\mathbf{z}}$ is the Wigner operator defined as

$$\hat{\Delta}_{\mathbf{z}} \doteq \frac{1}{2\pi} \int d^2\mathbf{a} e^{i\hat{\mathbf{a}} \wedge \mathbf{z}} \hat{T}_{\mathbf{a}}. \quad (2.24)$$

Since $\hat{\Delta}_{\mathbf{z}}$ is expressed through $\hat{T}_{\mathbf{a}}$, and $\hat{T}_{\mathbf{a}}$ is expressed through $\hat{\mathbf{z}}$, (2.23) also means that \hat{O} is expressible through $\hat{\mathbf{z}}$ and can be symbolically written as

$$\hat{O} = \mathfrak{D}(\hat{\mathbf{z}}) \equiv \mathfrak{D}(\hat{x}, \hat{k}). \quad (2.25)$$

In general, (2.25) is just a shorthand for (2.23). If \mathfrak{D} is representable as a sum $\mathfrak{D}_x(\hat{x}) + \mathfrak{D}_k(\hat{k})$, then \mathfrak{D} is identical to the operator symbol, O , except it is evaluated not on x and k but on the corresponding operators. The x - and k -representations of (2.25) are

$$\hat{O}_{\hat{x}} = \mathfrak{D}(x, -i\partial_x), \quad \hat{O}_{\hat{k}} = \mathfrak{D}(i\partial_k, k). \quad (2.26)$$

¹⁰Keep in mind, though, that \cdot is not a scalar product *per se* and does not define a length. In particular, the expression $\mathbf{z} \cdot \mathbf{z}$ is generally meaningless, because, by definition, $\mathbf{z} \cdot \mathbf{z} = x^2 + k^2$, yet x^2 and k^2 generally have non-matching dimensions.

¹¹See, for example, de Gosson (2006) and Tracy *et al.* (2014) for more details or Dodin *et al.* (2019, Supplemental Material) for a short summary with the key derivations.

¹²More general definitions of symbols exist (Martinez 2002) but will not be considered here.

For differential operators, the corresponding \mathfrak{D} are polynomials in $-i\partial_x$ (or in $i\partial_k$, depending on the context) and are typically known explicitly in a given problem. More generally, (2.26) can be understood as a symbolic pseudo-differential representation of \hat{Q} .

Equation (2.23) can also be represented as

$$\hat{O} = \frac{1}{2\pi} \int dx dk ds |\mathbf{e}_{\hat{x}}(x + s/2)\rangle O(x, k) e^{iks} \langle \mathbf{e}_{\hat{x}}(x - s/2)| \quad (2.27a)$$

$$= \frac{1}{2\pi} \int dx dk dr |\mathbf{e}_{\hat{k}}(k + r/2)\rangle O(x, k) e^{-irx} \langle \mathbf{e}_{\hat{k}}(k - r/2)|, \quad (2.27b)$$

and the symbol can be expressed as follows:

$$O(x, k) = \int ds \langle \mathbf{e}_{\hat{x}}(x + s/2)| \hat{O} | \mathbf{e}_{\hat{x}}(x - s/2)\rangle e^{-iks} \quad (2.28a)$$

$$= \int dr \langle \mathbf{e}_{\hat{k}}(k + r/2)| \hat{O} | \mathbf{e}_{\hat{k}}(k - r/2)\rangle e^{irx}. \quad (2.28b)$$

Note that the symbol of an operator is related to its x -representation via the Fourier transform with respect to the argument difference $x - x'$ at fixed $(x + x')/2$:

$$\langle \mathbf{e}_{\hat{x}}(x)| \hat{O} | \mathbf{e}_{\hat{x}}(x')\rangle = \frac{1}{2\pi} \int dk e^{ik(x-x')} O\left(\frac{x+x'}{2}, k\right), \quad (2.29)$$

and similarly for the k -representation. This also leads to the following notable properties of Weyl symbols:

$$\langle \mathbf{e}_{\hat{x}}(x)| \hat{O} | \mathbf{e}_{\hat{x}}(x)\rangle = \frac{1}{2\pi} \int dk O(x, k), \quad (2.30a)$$

$$\langle \mathbf{e}_{\hat{k}}(k)| \hat{O} | \mathbf{e}_{\hat{k}}(k)\rangle = \frac{1}{2\pi} \int dx O(x, k). \quad (2.30b)$$

An operator unambiguously determines its symbol and *vice versa*. We denote this isomorphism as $\hat{O} \Leftrightarrow O$. The mapping $\hat{O} \Rightarrow O$ is called the Wigner transform and $O \Rightarrow \hat{O}$ is called the Weyl transform. For uniformity, we will refer to these as the direct and inverse Wigner–Weyl transform. The Wigner–Weyl isomorphism \Leftrightarrow is natural in that it has the following properties:

$$\hat{1} \Leftrightarrow 1, \quad \hat{x} \Leftrightarrow x, \quad \hat{k} \Leftrightarrow k, \quad h(\hat{x}) \Leftrightarrow h(x), \quad h(\hat{k}) \Leftrightarrow h(k), \quad \hat{O}^\dagger \Leftrightarrow O^*, \quad (2.31)$$

where h is any function and \hat{O} is any operator. For multi-component operators, the Wigner–Weyl transform is applied element-by-element. In particular, if \hat{O} is a matrix operator, one finds that $\hat{O}^\dagger \Leftrightarrow O^\dagger$, where O is the symbol matrix.

The product of two operators maps to the so-called Moyal product, or star product, of their symbols:

$$\hat{A}\hat{B} \Leftrightarrow A(x, k) \star B(x, k) \doteq A(x, k) e^{i\hat{L}/2} B(x, k). \quad (2.32)$$

Here, $\hat{\mathcal{L}} \doteq \overleftarrow{\partial}_x \overrightarrow{\partial}_k - \overleftarrow{\partial}_k \overrightarrow{\partial}_x$ and the arrows indicate the directions in which the derivatives act. For example, $A \hat{\mathcal{L}} B$ is just the canonical Poisson bracket:

$$A \hat{\mathcal{L}} B = \{A, B\} \doteq \frac{\partial A}{\partial x} \frac{\partial B}{\partial k} - \frac{\partial A}{\partial k} \frac{\partial B}{\partial x} \equiv \frac{\partial A}{\partial z^\alpha} J^{\alpha\beta} \frac{\partial B}{\partial z^\beta}. \quad (2.33)$$

These formulae readily yield

$$h(\hat{x})\hat{k} \Leftrightarrow kh(x) + \frac{i}{2} h'(x), \quad \hat{k}h(\hat{x}) \Leftrightarrow kh(x) - \frac{i}{2} h'(x) \quad (2.34)$$

(the prime denotes the derivative with respect to the argument), so

$$kh(x) \Leftrightarrow \frac{1}{2}(h(\hat{x})\hat{k} + \hat{k}h(\hat{x})) = h(\hat{x})\hat{k} + \frac{1}{2}[\hat{k}, h(\hat{x})] = h(\hat{x})\hat{k} - \frac{i}{2} h'(\hat{x}), \quad (2.35)$$

or, in the x -representation,

$$kh(x) \Leftrightarrow -ih(x)\partial_x - ih'(x)/2. \quad (2.36)$$

Also, $h(\hat{k})e^{iK\hat{x}} \Leftrightarrow h(k) \star e^{iKx} = h(k + K/2)e^{iKx}$, and so on. In particular, the symbol of the commutator $[\hat{A}, \hat{B}]$ can be expressed through the so-called Moyal bracket $\{\{A, B\}\}$:

$$[\hat{A}, \hat{B}] \Leftrightarrow A \star B - B \star A \equiv i\{\{A, B\}\}. \quad (2.37)$$

The Moyal product and the Moyal bracket are particularly handy when $\partial_x \partial_k \sim \epsilon \ll 1$. (Sometimes, ϵ is equated with the GO parameter λ/L , but note that \star is symplectically invariant and the GO parameter is not. One might notice, though, that this ϵ is similar to the MGO parameter; cf. § 5.4.) Since $\hat{\mathcal{L}} = \mathcal{O}(\epsilon)$, one can express the Moyal product as an asymptotic series in powers of ϵ :

$$\star = \hat{1} + i\hat{\mathcal{L}}/2 - \hat{\mathcal{L}}^2/8 + \mathcal{O}(\epsilon^3). \quad (2.38)$$

Accordingly, the Moyal bracket can be approximated with the Poisson bracket:

$$\{\{A, B\}\} = \{A, B\} + \mathcal{O}(\epsilon^3) \quad (2.39)$$

(assuming that A and B are scalars or commuting matrices). Likewise, the symbol of any operator in the limit $\epsilon \rightarrow 0$ can be obtained simply by replacing \hat{x} with x and \hat{k} with k .

2.3. Wigner functions

Of particular interest in the Weyl symbol calculus are ‘density operators’

$$\hat{W}^{(\psi)} \doteq (2\pi)^{-1} |\psi\rangle \langle\psi| \quad (2.40)$$

for various $|\psi\rangle$. (The factor 2π is sometimes omitted, depending on a convention.) The symbol of such an operator, called the Wigner function,

$$W^{(\psi)} \doteq \text{symb} \hat{W}^{(\psi)}, \quad (2.41)$$

is a real function and can be expressed as follows:

$$W^{(\psi)}(x, k) = \frac{1}{2\pi} \int ds \psi_{\hat{x}}(x + s/2) \psi_{\hat{x}}^*(x - s/2) e^{-iks} \quad (2.42a)$$

$$= \frac{1}{2\pi} \int dr \psi_{\hat{k}}(k + r/2) \psi_{\hat{k}}^*(k - r/2) e^{irx}. \quad (2.42b)$$

(Like in the previous formulae, 2π here must be replaced with $(2\pi)^n$ if x is n -dimensional.) In particular, note the two special cases:

$$W^{(\psi)}(x, k) = \frac{1}{2\pi} \begin{cases} \delta(x - x_0), & |\psi\rangle = |\mathbf{e}_{\hat{x}}(x_0)\rangle, \\ \delta(k - k_0), & |\psi\rangle = |\mathbf{e}_{\hat{k}}(k_0)\rangle. \end{cases} \quad (2.43)$$

Any function bilinear in $\psi_{\hat{x}}$ and $\psi_{\hat{x}}^*$ can be expressed through $W^{(\psi)}$. For example, for any operators \hat{L} and \hat{R} , one has

$$(\hat{L}\psi_{\hat{x}}(x))(\hat{R}\psi_{\hat{x}}(x))^* = \int dk L(x, k) \star W^{(\psi)}(x, k) \star R^*(x, k), \quad (2.44)$$

where L and R are the corresponding symbols. As a corollary, one has

$$|\psi_{\hat{x}}(x)|^2 = \int dk W^{(\psi)}(x, k), \quad |\psi_{\hat{k}}(k)|^2 = \int dx W^{(\psi)}(x, k). \quad (2.45)$$

If $|\psi_{\hat{x}}(x)|^2$ and $|\psi_{\hat{k}}(k)|^2$ are interpreted as the densities of quanta in the x -space and the k -space, respectively, then one can attribute $W^{(\psi)}$ as a quasiprobability distribution of wave quanta in phase space. The prefix ‘quasi’ is commonly added because $W^{(\psi)}$ can be negative, which is not something that one expects from a probability distribution. That said, a Wigner function averaged over a sufficiently large phase-space volume $\Delta x \Delta k \gtrsim 1$ is guaranteed to be non-negative. Such a function can be understood as the spectrum of the two-point correlation function $\langle \psi_{\hat{x}}(x + s/2) \psi_{\hat{x}}^*(x - s/2) \rangle$ over s at given x and represents a local property of the field. For further details, see, for example, Dodin (2022).

3. Metaplectic transform

Now that we have introduced the Weyl symbol calculus and Wigner functions, let us expand on the MT definition introduced in § 1.2. What follows is not intended as a comprehensive theory of MTs (for that, one is referred to, for example, Littlejohn (1986), Martinez (2002) and Tracy *et al.* (2014), and papers cited therein) and we adhere to the accepted mathematical terminology only to the extent that it suits our purposes. If it helps, any parallels with the already existing theory of MTs can be considered accidental.

3.1. Definition

Consider a transformation of the base operators to some new base operators:

$$\hat{z} \equiv \begin{pmatrix} \hat{x} \\ \hat{k} \end{pmatrix} \mapsto \begin{pmatrix} \hat{q} \\ \hat{p} \end{pmatrix} \equiv \hat{y}. \quad (3.1)$$

Here, \hat{q} can be defined as any Hermitian operator whose eigenvalue space is \mathbb{R} and \hat{p} is defined such that its q -representation is $\hat{p}_{\hat{q}} = -i\partial_q$. Accordingly,

$$[\hat{q}, \hat{p}] = i, \quad (3.2)$$

and thus the corresponding symbols $q(x, k)$ and $p(x, k)$ satisfy¹³

$$\{q, p\} = 1. \quad (3.3)$$

In the GO limit, this corresponds to a variable transformation

$$\mathbf{z} \equiv \begin{pmatrix} x \\ k \end{pmatrix} \mapsto \begin{pmatrix} q \\ p \end{pmatrix} \equiv \mathbf{y} \quad (3.4)$$

with the following Jacobian matrix:

$$\mathcal{E} \doteq \begin{pmatrix} \partial_x q & \partial_k q \\ \partial_x p & \partial_k p \end{pmatrix}. \quad (3.5)$$

Since $\mathcal{E} \mathbf{J} \mathcal{E}^\top = \{q, p\} \mathbf{J}$ and $\{q, p\} \rightarrow 1$ by the GO limit of (3.3), this transformation conserves the symplectic form and therefore is canonical. Thus, (3.1) can be interpreted as a canonical transformation for operators.¹⁴ Also, as a reminder, canonical transformations conserve the Poisson bracket for any A and B ; i.e.

$$\{A, B\} \equiv \frac{\partial A}{\partial \mathbf{z}^\alpha} \mathbf{J}^{\alpha\beta} \frac{\partial B}{\partial \mathbf{z}^\beta} = \frac{\partial A}{\partial \mathbf{y}^\alpha} \mathbf{J}^{\alpha\beta} \frac{\partial B}{\partial \mathbf{y}^\beta}. \quad (3.6)$$

The operator transformation (3.1) induces the following transformation of fields:

$$\langle \mathbf{e}_{\hat{q}}(q) | \psi \rangle = \int dx \langle \mathbf{e}_{\hat{q}}(q) | \mathbf{e}_{\hat{x}}(x) \rangle \langle \mathbf{e}_{\hat{x}}(x) | \psi \rangle, \quad (3.7)$$

or, equivalently,

$$\psi_{\hat{q}}(q) = \int dx M(q, x) \psi_{\hat{x}}(x), \quad (3.8)$$

where we have introduced

$$M(q, x) \equiv M_{\hat{q} \leftarrow \hat{x}}(q, x) \doteq \langle \mathbf{e}_{\hat{q}}(q) | \mathbf{e}_{\hat{x}}(x) \rangle. \quad (3.9)$$

When $\hat{q} = \hat{k}$ (and thus $\hat{p} = -\hat{x}$), the metaplectic transform is simply the Fourier transform, as seen by comparing (3.7) with (2.18). Since the function $M_{\hat{q} \leftarrow \hat{x}}(q, x)$ defines the corresponding operator $\hat{M}_{\hat{q} \leftarrow \hat{x}}$, one can as well express (3.8) as¹⁵

$$\psi_{\hat{q}} = \hat{M} \psi_{\hat{x}}, \quad \hat{M} \equiv \hat{M}_{\hat{q} \leftarrow \hat{x}}. \quad (3.10)$$

We call this operator the MT induced by the operator transformation (3.1).¹⁶ Its inverse is the metaplectic operator $(\hat{M}_{\hat{q} \leftarrow \hat{x}})^{-1} = \hat{M}_{\hat{x} \leftarrow \hat{q}} \equiv \hat{M}$ induced by the inverse transformation $(\hat{q}, \hat{p}) \mapsto (\hat{x}, \hat{k})$ and has the kernel

$$M_{\hat{x} \leftarrow \hat{q}}(x, q) \equiv \bar{M}(x, q) = \langle \mathbf{e}_{\hat{x}}(x) | \mathbf{e}_{\hat{q}}(q) \rangle = M_{\hat{q} \leftarrow \hat{x}}^*(q, x). \quad (3.11)$$

By (2.10), this means that $(\hat{M}_{\hat{q} \leftarrow \hat{x}})^{-1} = \hat{M}_{\hat{q} \leftarrow \hat{x}}^\dagger$, i.e. \hat{M} is unitary.

¹³Using the terminology to be introduced §3.4, $\mathbf{y}(\mathbf{z}) \equiv (q(x, k), p(x, k))^\top$ is a \mathbf{z} -symbol of \mathbf{y} , to be denoted \mathbf{y}_z . Likewise, $\mathbf{z}(\mathbf{y}) \equiv (x(q, p), k(q, p))^\top$ is a \mathbf{y} -symbol of \mathbf{z} , to be denoted \mathbf{z}_y .

¹⁴Note that, beyond the GO limit, having $\{q, p\} = 1$ does not guarantee (3.2) and *vice versa*.

¹⁵As a special case, (3.4) can represent the evolution of a classical Hamiltonian system, with $(x, k) \equiv (x_0, k_0)$ as the initial coordinates and $(q, p) \equiv (x_T, k_T)$ as the coordinates at time T . Then, the MT is the propagator for the wavefunction of the corresponding quantum system.

¹⁶Some authors use other approaches to defining MTs, particularly, MTs corresponding to specific variable transformations. It is not our goal to overview this subject here, but see, for example, Mello & Moshinsky (1975) and Littlejohn (1986), and the references therein.

A sequence of MTs $\hat{M}_1 \hat{M}_0$ corresponding to the operator transformations $\hat{x}_0 \mapsto \hat{x}_1 \mapsto \hat{x}_2$ is an MT \hat{M}_2 corresponding to the operator transformation $\hat{x}_0 \mapsto \hat{x}_2$, because

$$\begin{aligned} \psi_{\hat{x}_2}(x_2) &= \int dx_1 M_1(x_2, x_1) \psi_{\hat{x}_1}(x_1) \\ &= \int dx_1 dx_0 M_1(x_2, x_1) M_0(x_1, x_0) \psi_{\hat{x}_0}(x_0) \\ &= \int dx_0 \mathcal{M}(x_2, x_0) \psi_{\hat{x}_0}(x_0), \end{aligned} \quad (3.12)$$

where

$$\begin{aligned} \mathcal{M}(x_2, x_0) &\doteq \int dx_1 M_1(x_2, x_1) M_0(x_1, x_0) \\ &= \int dx_1 \langle \mathbf{e}_{\hat{x}_2}(x_2) | \mathbf{e}_{\hat{x}_1}(x_1) \rangle \langle \mathbf{e}_{\hat{x}_1}(x_1) | \mathbf{e}_{\hat{x}_0}(x_0) \rangle \\ &= \langle \mathbf{e}_{\hat{x}_2}(x_2) | \mathbf{e}_{\hat{x}_0}(x_0) \rangle \\ &= M_2(x_2, x_0). \end{aligned} \quad (3.13)$$

Thus, MTs form a group.

3.2. *M*-waves

At given x considered as a parameter, the function $M(q, x) \equiv \langle \mathbf{e}_{\hat{q}}(q) | \mathbf{e}_{\hat{x}}(x) \rangle$ can be understood as a field on the q -space, specifically, the q -representation of the vector $|\mathbf{e}_{\hat{x}}(x)\rangle \equiv |M_x\rangle$. As such, M can be acted upon by linear operators like any other field. Let us explore the effect of \hat{x} on it. By definition,

$$\hat{x} |\mathbf{e}_{\hat{x}}(x)\rangle = x |\mathbf{e}_{\hat{x}}(x)\rangle. \quad (3.14)$$

This leads to

$$(\hat{x}_{\hat{q}} M)(q, x) = x M(q, x), \quad (3.15)$$

where $(\hat{x}_{\hat{q}} M)(q, x) \equiv \langle \mathbf{e}_{\hat{q}}(q) | \hat{x} |M_x\rangle$. Let us symbolically represent $\hat{x}_{\hat{q}}$ through the new base operators $\hat{q}_{\hat{q}} = q$ and $\hat{p}_{\hat{q}} = -i\partial_q$ as $X(q, -i\partial_q)$.¹⁷ This leads to the following equation:

$$X(q, -i\partial_q) M(q, x) = x M(q, x). \quad (3.16)$$

Equation (3.16) is a (pseudo)differential equation for M as a function of q , with x as a parameter. Thus, the solution of (3.16) is defined up only to a function of x and another equation for M is needed to specify this function. To derive such an equation, note that $\bar{M}(x, q) \equiv \langle \mathbf{e}_{\hat{x}}(x) | \mathbf{e}_{\hat{q}}(q) \rangle \equiv M^*(q, x)$ can be understood as the x -representation of $|\mathbf{e}_{\hat{q}}(q)\rangle$. Hence, on the one hand,

$$(\hat{k}_{\hat{x}} \bar{M})(x, q) = -i\partial_x \bar{M}(x, q) = -i\partial_x M^*(q, x). \quad (3.17)$$

On the other hand,

$$(\hat{k}_{\hat{x}} \bar{M})(x, q) = \langle \mathbf{e}_{\hat{x}}(x) | \hat{k} | \mathbf{e}_{\hat{q}}(q) \rangle = \langle \mathbf{e}_{\hat{q}}(q) | \hat{k} | \mathbf{e}_{\hat{x}}(x) \rangle^* = (\hat{k}_{\hat{q}} M)^*(q, x). \quad (3.18)$$

¹⁷This notation is explained in § 2.2.

Thus, $(\hat{k}_{\hat{q}}M)(q, x) = i\partial_x M(q, x)$, or, assuming the notation $\hat{k}_{\hat{q}} = K(q, -i\partial_q)$,

$$K(q, -i\partial_q)M(q, x) = i\partial_x M(q, x). \quad (3.19)$$

Equation (3.19) can be interpreted as a Schrödinger equation (SE) for M , with x serving as the time variable. In this sense, M is a wave, which we call an M -wave. Likewise, in retrospect, (3.16) can be interpreted as an SE for a stationary wave with ‘energy’ x .

Together, (3.16) and (3.19) define M up to a constant complex factor. From the normalisation condition, one finds that

$$\begin{aligned} \delta(x' - x'') &= \langle \mathbf{e}_{\hat{x}}(x') | \mathbf{e}_{\hat{x}}(x'') \rangle \\ &= \int dq \langle \mathbf{e}_{\hat{x}}(x') | \mathbf{e}_{\hat{q}}(q) \rangle \langle \mathbf{e}_{\hat{q}}(q) | \mathbf{e}_{\hat{x}}(x'') \rangle \\ &= \int dq M^*(q, x') M(q, x''). \end{aligned} \quad (3.20)$$

This means that, for any given x , M satisfies

$$\int dq ds M^*(q, x + s/2) M(q, x - s/2) = 1. \quad (3.21)$$

(This is also seen from the definition of M as a representation of a normalised eigenvector.) This defines the normalisation factor up to a constant phase, which can be anything depending how the eigenvectors of \hat{q} are chosen relative to those of \hat{x} . This choice constitutes a gauge freedom of the theory.

Similar equations apply to the inverse-MT kernel $\bar{M}(x, q)$. That is, assuming the notation $\hat{q}_{\hat{x}} = Q(x, -i\partial_x)$ and $\hat{p}_{\hat{x}} = P(x, -i\partial_x)$, one has

$$Q(x, -i\partial_x)\bar{M}(x, q) = q\bar{M}(x, q), \quad (3.22)$$

$$P(x, -i\partial_x)\bar{M}(x, q) = i\partial_q\bar{M}(x, q). \quad (3.23)$$

Equation (3.22) can be interpreted as an SE for a wave \bar{M} , with q serving as the time variable. Because $\bar{M}(x, q) \equiv M^*(x, q)$, we will also call it an M -wave. Also, (3.22) can be interpreted as an SE for a stationary wave with ‘energy’ q . Like before, one also finds

$$\begin{aligned} \delta(q' - q'') &= \langle \mathbf{e}_{\hat{q}}(q') | \mathbf{e}_{\hat{q}}(q'') \rangle \\ &= \int dx \langle \mathbf{e}_{\hat{q}}(q') | \mathbf{e}_{\hat{x}}(x) \rangle \langle \mathbf{e}_{\hat{x}}(x) | \mathbf{e}_{\hat{q}}(q'') \rangle \\ &= \int dx \bar{M}^*(x, q') \bar{M}(x, q''), \end{aligned} \quad (3.24)$$

whence

$$\int dx ds \bar{M}^*(x, q + s/2) \bar{M}(x, q - s/2) = 1. \quad (3.25)$$

Let us summarise our intermediate results. M -waves are normalised solutions of SEs (3.19) and (3.23), which can also be written in the form (4.1):

$$(K(q, -i\partial_q) - i\partial_x)M = 0, \quad (P(x, -i\partial_x) - i\partial_q)\bar{M} = 0. \quad (3.26)$$

Their solutions can be formally expressed as follows:

$$M(q, x) = \exp(-ix\hat{k}_{\hat{q}})M(q, 0), \quad \bar{M}(x, q) = \exp(-iq\hat{p}_{\hat{x}})\bar{M}(x, 0). \quad (3.27)$$

It is generally impossible to find explicit analytic expressions for the propagators $\exp(-ix\hat{k}_{\hat{q}})$ and $\exp(-iq\hat{p}_{\hat{x}})$ for a given variable transformation. However, it is possible to do so for some important special cases and also asymptotically. This will be discussed later, in §§ 3.5 and 4.3.

3.3. Symplectic pseudo-measure

As a function of two variables, an M -wave can be considered as a field on the two-dimensional ‘timespace’ (x, q) . The variable x serves as the time variable in the Schrödinger equation (3.19), so $i\partial_x$ serves as the frequency operator and $-k$ serves as the frequency variable. Hence, MGO naturally involves Wigner functions of M in the form

$$W_z^{(M)}(x, q, -k, p) \equiv \mu(\mathbf{y}, \mathbf{z})/2\pi \quad (3.28)$$

(the utility of the added coefficient 2π will become clear shortly), i.e.

$$\mu(\mathbf{y}, \mathbf{z}) = \frac{1}{2\pi} \int ds ds' M(q + s'/2, x + s/2) M^*(q - s'/2, x - s/2) e^{-ips' + iks}. \quad (3.29)$$

For the inverse transformation, the corresponding function is

$$\begin{aligned} \bar{\mu}(\mathbf{z}, \mathbf{y}) &= \frac{1}{2\pi} \int ds ds' \bar{M}(x + s/2, q + s'/2) \bar{M}^*(x - s/2, q - s'/2) e^{-iks + ips'} \\ &= \frac{1}{2\pi} \int ds ds' M^*(q + s'/2, x + s/2) M(q - s'/2, x - s/2) e^{ips' - iks} \\ &= \mu^*(\mathbf{y}, \mathbf{z}). \end{aligned} \quad (3.30)$$

Like any Wigner function, μ is real, so one can also express (3.30) simply as

$$\bar{\mu}(\mathbf{z}, \mathbf{y}) = \mu(\mathbf{y}, \mathbf{z}). \quad (3.31)$$

Also note that

$$\begin{aligned} \int \mu(\mathbf{y}, \mathbf{z}) d\mathbf{z} &= \int dx ds ds' M(q + s'/2, x + s/2) M^*(q - s'/2, x - s/2) \delta(s) e^{-ips'} \\ &= \int ds' e^{-ips'} \int dx M(q + s'/2, x) M^*(q - s'/2, x) \\ &= \int ds' e^{-ips'} \delta(s') \\ &= 1 \end{aligned} \quad (3.32)$$

(here we used (3.24)) and, similarly, $\int \mu(\mathbf{y}, \mathbf{z}) d\mathbf{y} = 1$ too. In summary then,

$$\int \mu(\mathbf{y}, \mathbf{z}) d\mathbf{z} = \int \mu(\mathbf{y}, \mathbf{z}) d\mathbf{y} = 1. \quad (3.33)$$

As to be seen later (§ 3.4), these serve as pseudo-measures on the \mathbf{z} -space and the \mathbf{y} -space, respectively. (The prefix ‘pseudo’ is added because μ can be negative.) From here, the arguments of μ will be occasionally omitted where they are obvious from the context.

3.4. Transformation of operators

A metaplectic transform induces a transformation of the operator representation $\hat{O}_{\hat{x}} \mapsto \hat{O}_{\hat{q}}$, which is found as follows. Note that

$$(\hat{O}_{\hat{q}}\psi_{\hat{q}})(q) = \langle \mathbf{e}_{\hat{q}}(q) | \hat{O} | \psi \rangle = \int dq' \langle \mathbf{e}_{\hat{q}}(q) | \hat{O} | \mathbf{e}_{\hat{q}}(q') \rangle \langle \mathbf{e}_{\hat{q}}(q') | \psi \rangle. \quad (3.34)$$

Using (2.9), one can also represent this as follows:

$$\begin{aligned} (\hat{O}_{\hat{q}}\psi_{\hat{q}})(q) &= \int dq' dx' dx'' \langle \mathbf{e}_{\hat{q}}(q) | \mathbf{e}_{\hat{x}}(x') \rangle O_{\hat{x}}(x', x'') \langle \mathbf{e}_{\hat{x}}(x'') | \mathbf{e}_{\hat{q}}(q') \rangle \psi_{\hat{q}}(q') \\ &= \int dq' dx' dx'' M(q, x') O_{\hat{x}}(x', x'') M^*(q', x'') \psi_{\hat{q}}(q') \\ &= \int dq' (\hat{M} \hat{O}_{\hat{x}} \hat{M}^\dagger)(q, q') \psi_{\hat{q}}(q'), \end{aligned} \quad (3.35)$$

whence

$$\hat{O}_{\hat{q}} = \hat{M} \hat{O}_{\hat{x}} \hat{M}^\dagger. \quad (3.36)$$

Let us use the following formula for the Weyl symbol of a given operator (cf. (2.28a)):

$$O_y(q, p) = \int ds' \langle \mathbf{e}_{\hat{q}}(q + s'/2) | \hat{O} | \mathbf{e}_{\hat{q}}(q - s'/2) \rangle e^{-ips'}. \quad (3.37)$$

The subindex in the symbol's notation refers to the fact that this symbol is defined on the phase space $\mathbf{y} \doteq (q, p)^\top$. We will call it the \mathbf{y} -symbol of \hat{O} , and the symbol defined on the phase space $\mathbf{z} \doteq (x, k)^\top$ will be called the \mathbf{z} -symbol of \hat{O} . Recall that

$$\hat{O} = \frac{1}{2\pi} \int dx dk ds |\mathbf{e}_{\hat{x}}(x + s/2)\rangle O_z(x, k) \langle \mathbf{e}_{\hat{x}}(x - s/2) | e^{iks}. \quad (3.38)$$

Then,

$$\begin{aligned} O_y(q, p) &= \frac{1}{2\pi} \int dx dk ds ds' \langle \mathbf{e}_{\hat{q}}(q + s'/2) | \mathbf{e}_{\hat{x}}(x + s/2) \rangle \langle \mathbf{e}_{\hat{x}}(x - s/2) | \mathbf{e}_{\hat{q}}(q - s'/2) \rangle \\ &\quad \times O_z(x, k) e^{-ips' + iks}. \end{aligned} \quad (3.39)$$

Expressing the brackets through M , one obtains the following expression for the \mathbf{y} -symbol in terms of the \mathbf{z} -symbol:

$$\begin{aligned} O_y(q, p) &= \frac{1}{2\pi} \int dx dk ds ds' M(q + s'/2, x + s/2) M^*(q - s'/2, x - s/2) \\ &\quad \times O_z(x, k) e^{-ips' + iks}, \end{aligned} \quad (3.40)$$

and the inverse transformation can be written similarly. Equivalently, this result can be expressed in the following compact form:

$$O_y(\mathbf{y}) = \int O_z(\mathbf{z}) \mu(\mathbf{y}, \mathbf{z}) d\mathbf{z}, \quad O_z(\mathbf{z}) = \int O_y(\mathbf{y}) \mu(\mathbf{y}, \mathbf{z}) d\mathbf{y}. \quad (3.41)$$

(If $O_z \equiv 1$, then (3.40) readily yields $O_y \equiv 1$, and *vice versa*, by (3.33).) In particular, these allow one to recalculate the Wigner functions from one representation to another.

3.5. Examples

In this section, we present several examples of MTs that are used as building blocks of the theory presented in later sections.

3.5.1. Shift

For example, let us consider a variable transformation that is a phase-space shift by a constant vector $\Delta \doteq (\Delta_q, \Delta_p)^\top$:

$$\hat{q} = \hat{x} + \Delta_q \hat{1}, \quad \hat{p} = \hat{k} + \Delta_p \hat{1}. \quad (3.42)$$

This corresponds to

$$X(q, -i\partial_q) = q - \Delta_q, \quad K(q, -i\partial_q) = -i\partial_q - \Delta_p. \quad (3.43)$$

Then, (3.16) requires that $M(q, x) = C(q)\delta(q - x - \Delta_q)$, where C is some function. From (3.19), one finds that $(-i\partial_q - i\partial_x - \Delta_p)M(q, x) = 0$, whence $\partial_q C = i\Delta_p C$. With (3.21) for the normalisation, this gives

$$M(q, p) = e^{i\Delta_p q} \delta(q - x - \Delta_q) \quad (3.44)$$

(up to an arbitrary constant phase factor, which we choose equal to one) and, therefore,

$$\psi_{\hat{q}}(q) = e^{i\Delta_p q} \psi_{\hat{x}}(q - \Delta_q). \quad (3.45)$$

Notice that even at $\Delta_q = 0$, when q is the same as x , the function $\psi_{\hat{q}}$ is not the same as $\psi_{\hat{x}}$ due to non-zero Δ_p . Also,

$$\begin{aligned} \mu &= \frac{1}{2\pi} \int ds ds' e^{i\Delta_p s' - ips' + iks} \delta(q + s'/2 - x - s/2 - \Delta_q) \delta(q - s'/2 - x + s/2 - \Delta_q) \\ &= \frac{1}{2\pi} \int ds ds' e^{i(\Delta_p - p)s' + iks} \delta(q - x - \Delta_q) \delta(s - s') \\ &= \frac{1}{2\pi} \int ds e^{i(k + \Delta_p - p)s} \delta(q - x - \Delta_q) \\ &= \delta(x - q + \Delta_q) \delta(k - p + \Delta_p). \end{aligned} \quad (3.46)$$

Then, by (3.41), one obtains

$$O_y(q, p) = O_z(q - \Delta_q, p - \Delta_p) = O_z(x(q, p), k(q, p)). \quad (3.47)$$

Simply put, this means that $O_y(\mathbf{y}) = O_z(\mathbf{z})$; i.e. the symbol of an operator is invariant with respect to phase-space shifts.

3.5.2. Rescaling

Let us consider a rescaling canonical transformation:

$$\hat{q} = \hat{x}/\alpha, \quad \hat{p} = \alpha\hat{k}, \quad (3.48)$$

where α is a non-zero real constant. (Both equations contain the same factor α to keep the transformation canonical.) Clearly, $M(q, x) = C\delta(q - x/\alpha)$, where C is a constant. From (3.21), one finds that $C = |\alpha|^{-1/2}$ up to an arbitrary phase, so

$$M(q, x) = |\alpha|^{-1/2} \delta(q - x/\alpha), \quad \psi_{\hat{q}}(q) = |\alpha|^{1/2} \psi_{\hat{x}}(\alpha q). \quad (3.49)$$

3.5.3. Linear symplectic transformation

Let us also consider a linear symplectic transformation (LST), that is,

$$\begin{pmatrix} \hat{q} \\ \hat{p} \end{pmatrix} = \underbrace{\begin{pmatrix} A & B \\ C & D \end{pmatrix}}_{\mathbf{S}} \begin{pmatrix} \hat{x} \\ \hat{k} \end{pmatrix}, \quad (3.50)$$

where the (constant) coefficients satisfy

$$AD - BC = 1, \quad (3.51)$$

i.e. $\det \mathbf{S} = 1$.¹⁸ This subsumes relabelling coordinates and momenta; specifically, $A = D = 0$ and $B = -C = \pm 1$ corresponds to the transformation $(\hat{q}, \hat{p}) = (\pm \hat{k}, \mp \hat{x})$. (Note that \hat{k} and \hat{x} necessarily enter the latter formula with opposite signs to ensure symplecticity.) The rescaling transformation considered in §3.5.2 is also subsumed as a limit ($B \rightarrow 0$, $C \rightarrow 0$ and $A \rightarrow 1/D$).

Equivalently, (3.50) can be represented in the following inverted form:

$$\begin{pmatrix} \hat{x} \\ \hat{k} \end{pmatrix} = \underbrace{\begin{pmatrix} \bar{A} & \bar{B} \\ \bar{C} & \bar{D} \end{pmatrix}}_{\bar{\mathbf{S}}} \begin{pmatrix} \hat{q} \\ \hat{p} \end{pmatrix}, \quad (3.52)$$

where $\bar{\mathbf{S}} = \mathbf{S}^{-1}$ is also symplectic; specifically,

$$\begin{pmatrix} \bar{A} & \bar{B} \\ \bar{C} & \bar{D} \end{pmatrix} = \begin{pmatrix} D & -B \\ -C & A \end{pmatrix}, \quad (3.53)$$

so $\det \bar{\mathbf{S}} = \det \mathbf{S} = 1$. Then, (3.16) and (3.19) yields

$$(\bar{A}q - i\bar{B}\partial_q - x)M(q, x) = 0, \quad (\bar{C}q - i\bar{D}\partial_q - i\partial_x)M(q, x) = 0. \quad (3.54)$$

With the normalisation (3.21) taken into account, a straightforward calculation yields the well-known result (Littlejohn 1986):

$$M(q, x) = (2\pi i \bar{B})^{-1/2} \exp\left(-\frac{i(\bar{A}q^2 - 2xq + \bar{D}x^2)}{2\bar{B}}\right). \quad (3.55)$$

As usual, an arbitrary constant phase can be added to the overall expression, depending on a convention. The phase of the square root as a function of \bar{B} must be chosen such that MTs remain a group. Loosely speaking, one can require that a phase-space rotation¹⁹ by $\pm 2\pi$ causes the phase of \bar{B} to change by $\pm 2\pi$ too; then, the phase of $\bar{B}^{1/2}$ changes by $\pm \pi$. (For a detailed discussion, see Littlejohn (1986).) Then, the MT corresponding to a complete phase-space rotation in either direction is not an identity operator $\hat{1}$ but $-\hat{1}$.

¹⁸Additional constraints apply at $\dim x > 1$, when A , B , C and D are matrices. See, for example, Lopez & Dodin (2019).

¹⁹Formally, a rotation matrix corresponds to $\bar{A} = \bar{D} = \cos \alpha$ and $-\bar{C} = \bar{B} = \sin \alpha$; hence, the rotation angle is usually defined as $\alpha = \arg(\bar{A} + i\bar{B})$. Remember, though, that \bar{A} and \bar{B} generally have different units (or, otherwise, are normalised arbitrarily). Thus, this expression for α is questionable unless \bar{A} or \bar{B} is zero, i.e. unless α is $\pi/2$ times an integer irrespective of the units.

The corresponding transformation of symbols is calculated as follows. First, note that

$$\begin{aligned} & M(q + s'/2, x + s/2) \bar{M}^*(q - s'/2, x - s/2) e^{-ips' + iks} \\ &= \frac{1}{2\pi|\bar{B}|} \exp\left(\frac{is(q - \bar{D}x + \bar{B}k) + is'(x - \bar{A}q - \bar{B}p)}{\bar{B}}\right), \end{aligned} \quad (3.56)$$

and

$$\begin{aligned} \mu &= \int \frac{ds'}{2\pi|\bar{B}|} \exp\left(\frac{is'(x - \bar{A}q - \bar{B}p)}{\bar{B}}\right) \int \frac{ds}{2\pi} \exp\left(is\left(\frac{q - \bar{D}x}{\bar{B}} + k\right)\right) \\ &= \delta(x - \bar{A}q - \bar{B}p) \delta\left(k - \frac{\bar{D}x - q}{\bar{B}}\right) \\ &= \delta(x - \bar{A}q - \bar{B}p) \delta\left(k - \frac{\bar{D}(\bar{A}q + \bar{B}p) - q}{\bar{B}}\right) \\ &= \delta(x - \bar{A}q - \bar{B}p) \delta(k - \bar{C}q - \bar{D}p), \end{aligned} \quad (3.57)$$

where we used (3.51). Then, by (3.41), one obtains

$$O_y(q, p) = O_z(\bar{A}q + \bar{B}p, \bar{C}q + \bar{D}p) = O_z(x(q, p), k(q, p)). \quad (3.58)$$

This shows that the symbols of operators are invariant with respect to LSTs. In particular, Wigner functions are LST-invariant. Also, for $\hat{O} = \hat{y}$, (3.58) gives $z_y = \bar{S}y$ and, similarly, $y_z = Sz$; i.e. the symbols of the base operators transform at LSTs exactly as the base operators themselves.

3.5.4. Eikonal transform

Let us also consider the ‘eikonal transform’ (this term will become clear shortly):

$$\hat{q} = \hat{x}, \quad \hat{p} = \hat{k} - \mathcal{K}(\hat{x}) = \hat{k} - \mathcal{K}(\hat{q}), \quad (3.59)$$

which satisfies the commutation relation (3.2). Like in § 3.5.1, one finds that

$$M(q, p) = e^{-i\theta(q)} \delta(q - x), \quad (3.60)$$

where $\theta \doteq \int \mathcal{K}(q) dq = \int \mathcal{K}(x) dx$. Then, $\psi_{\hat{q}}$ is given by

$$\psi_{\hat{q}}(q) = \psi_{\hat{q}}(x) = e^{-i\theta(x)} \psi_{\hat{x}}(x), \quad (3.61)$$

and

$$\begin{aligned} \mu &= \frac{1}{2\pi} \int ds ds' e^{-i\theta(q+s'/2) + i\theta(q-s'/2) - ips' + iks} \delta(q + s'/2 - x - s/2) \delta(q - s'/2 - x + s/2) \\ &= \frac{1}{2\pi} \int ds ds' e^{-i\theta(q+s'/2) + i\theta(q-s'/2) - ips' + iks} \delta(s' - s) \delta(q - x) \\ &= \frac{1}{2\pi} \int ds e^{-i\theta(q+s/2) + i\theta(q-s/2) + i(k-p)s} \delta(q - x). \end{aligned} \quad (3.62)$$

Then, by (3.41), the operator symbols are transformed as follows:

$$O_y(q, p) = \frac{1}{2\pi} \int dk ds O_z(q, k) e^{i(k-p)s - i\theta(q+s/2) + i\theta(q-s/2)}. \quad (3.63)$$

Using Taylor expansion, one has

$$\theta\left(q + \frac{s}{2}\right) - \theta\left(q - \frac{s}{2}\right) = \mathcal{K}(q)s + \frac{1}{24} \frac{\partial^2 \mathcal{K}(q)}{\partial q^2} s^3 + \dots, \quad (3.64)$$

where we used $\partial_q \theta(q) = \mathcal{K}(q)$. Thus,

$$\begin{aligned} O_y(q, p) &= \frac{1}{2\pi} \int dk ds O_z(q, k) \left(1 - \frac{i}{24} \frac{\partial^2 \mathcal{K}(q)}{\partial q^2} s^3 + \dots\right) e^{i(k-p-\mathcal{K}(q))s} \\ &= \frac{1}{2\pi} \left(1 - \frac{1}{24} \frac{\partial^2 \mathcal{K}(q)}{\partial q^2} \frac{\partial^3}{\partial p^3} + \dots\right) \int dk ds O_z(q, k) e^{i(k-p-\mathcal{K}(q))s} \\ &= \frac{1}{2\pi} \int dk ds O_z(q, k) e^{i(k-p-\mathcal{K}(q))s} + \mathcal{O}(\epsilon^2) \\ &= O_z(q, p + \mathcal{K}(q)) + \mathcal{O}(\epsilon^2). \end{aligned} \quad (3.65)$$

In particular, if $\theta(q)$ is the eikonal and $\psi_{\hat{q}}$ is the wave envelope, then \mathcal{K} serves as the reference wavevector and the symbol transformation reproduces the one derived, for example, by Dodin *et al.* (2019). The importance of (3.65) in this case is in that, within GO, where $\mathcal{O}(\epsilon^2)$ is negligible, one may use

$$O_y(q, p) \approx O_z(q, p + \mathcal{K}(q)) = O_z(x(q, p), k(q, p)). \quad (3.66)$$

4. Geometrical optics through metaplectic transforms

4.1. Basic equations

The Weyl symbol calculus and MTs provide a natural framework for formulating GO equations (Tracy *et al.* 2014). Here, we restate the conventional GO (McDonald 1988; Dodin *et al.* 2019), which will be an important reference point for our further discussion. Let us start with a generic linear wave equation²⁰

$$\hat{D}_{\hat{x}} \psi_{\hat{x}} = 0. \quad (4.1)$$

Here, $\psi_{\hat{x}}$ is a field on some x -space (generally, timespace), which we assume one-dimensional for simplicity, and $\hat{D}_{\hat{x}}$ is some dispersion operator, which can be an integral operator or a differential operator as a special case. Suppose that $\psi_{\hat{x}}$ is quasimonochromatic, i.e. representable as

$$\psi_{\hat{x}}(x) = e^{i\theta(x)} \Psi(x), \quad (4.2)$$

where the (real) phase θ is fast compared with its gradient²¹ $k(x) \doteq \partial_x \theta(x)$ and to the envelope Ψ . The evolution of the latter can be described by the following equation:

$$\hat{D}_{\hat{x}} \Psi = 0, \quad \hat{D} \doteq e^{-i\theta(\hat{x})} \hat{D} e^{i\theta(\hat{x})}. \quad (4.3)$$

Our goal is to approximate the operator \hat{D} with a low-order differential operator.

²⁰Specific $\hat{D}_{\hat{x}}$ that govern electromagnetic waves in dielectric media, plasma in particular, are discussed, for example, by Dodin *et al.* (2019, Section VI).

²¹The notation $k(x)$ denotes a given function of x . It should not be confused with k as a variable, which is typed without an argument and represents a generic coordinate in the spectral space (as in, for example, (2.18)). This abuse of notation will be continued later and adopted for other variables as well.

As we have just discussed (§ 3.5.4), the mapping from $\psi_{\hat{x}}$ to Ψ can be considered as an MT that corresponds to the variable transformation (3.59), so

$$\hat{D}_{\hat{x}} = \hat{D}_{\hat{q}}, \quad \hat{x} = \hat{q}, \quad \hat{p} = \hat{k} - k(\hat{x}), \quad (4.4)$$

i.e. the envelope-evolution operator is just the q -representation of the original ‘full-wave’ dispersion operator \hat{D} . Then, its symbol can be calculated using (3.66). Since Ψ that $\hat{D}_{\hat{x}}$ acts upon is smooth, only small values of p matter in this case, so one can further approximate (3.66) with its first-order Taylor expansion:²²

$$D_y(q, p) \approx D_z(q, k(q)) + pV(q), \quad V(q) \doteq (\partial_p D_z(q, p))_{p=k(q)}. \quad (4.5)$$

Loosely, the expansion parameter here is the ratio of the characteristic wavenumber p of the envelope and $k(q)$. Denoting the former as the inverse inhomogeneity scale, $1/L$, and introducing the wavelength $\lambda \sim 1/k(q)$, this small parameter can be roughly estimated as

$$\epsilon \sim \lambda/L. \quad (4.6)$$

Using the analogue of (2.36) for base operators \hat{q} and \hat{p} , one immediately obtains

$$\hat{D}_{\hat{q}} \approx D_z(q, k(q)) - iV(q)\partial_q - \frac{i}{2} V'(q), \quad (4.7)$$

where the prime denotes the derivative with respect to the argument, so $V'(q) = (\partial_{qp}^2 D_z(q, p) + p'(q) \partial_{pp}^2 D_z(q, p))_{p=p(q)}$. Using (4.4), one can then rewrite (4.3) as follows:

$$(H_z(x) + i\Gamma_z(x))\Psi - iV(x)\partial_x\Psi - \frac{i}{2} V'(x)\Psi \approx 0, \quad (4.8)$$

where we introduced $H_z(x) \doteq \text{re}D_z(x, k(x))$ and $\Gamma_z(x) \doteq \text{im}D_z(x, k(x))$. (For vector waves, these would be the Hermitian and anti-Hermitian parts of $D_z(q, k(q))$, correspondingly.) It is common to assume that $\text{re}D_z$ is an order-one function, while $\text{im}D_z$ is order- ϵ . Also, remember that $\partial_x\Psi$ and V' are order- ϵ as well, so one can redefine V as

$$V(x) \doteq (\partial_k H_z(x, k))_{k=k(x)} \quad (4.9)$$

without loss of accuracy. Then, all coefficients in (4.8) are real, except i *per se*.

Note that the function θ has been unspecified so far. One can choose it such that²³

$$0 = H_z(\mathbf{z}) \equiv H_z(x, \partial_x\theta) \quad (4.10)$$

(here, $\mathbf{z} \equiv (x, k)^\top$, as usual), which can be considered as a local dispersion relation or a Hamilton–Jacobi equation for θ . For given initial conditions, this determines the ‘reference ray’ trajectory $k(x)$, which can also be described by Hamilton’s ray equations (Appendix A)

$$\dot{\mathbf{z}} = \mathbf{J}\partial_{\mathbf{z}}H_z \equiv \{\mathbf{z}, H_z\}. \quad (4.11)$$

²²The second-order Taylor expansion is used for modelling transverse diffraction of quasioptical beams in multiple dimensions. Assuming that the transverse scale of a beam is much smaller than its longitudinal scale, (3.66) remains applicable in that case (Dodin *et al.* 2019).

²³For vector waves, one might want to adopt a slightly different equation for $k(x)$. We revisit this subject in a broader context in § 7.

The dot in \dot{z} denotes a derivative with respect to the ray time (which can be different from the physical time t ; see [Appendix A](#)) and H_z serves as the ray Hamiltonian. Accordingly, we henceforth address \hat{H} , i.e. the Hermitian part of \hat{D} , as the wave Hamiltonian and \hat{F} , i.e. the anti-Hermitian part of \hat{D} , as the dissipation operator.

The reference ray determines the coefficients in (4.8), which becomes

$$V(x)\partial_x\Psi + \frac{1}{2}V'(x)\Psi = \Gamma_z(x)\Psi, \quad \partial_x(V(x)|\Psi|^2) = 2\Gamma_z(x)|\Psi|^2. \quad (4.12)$$

(As a reminder, $V'(x)$ in these equations is generally not just $\partial_{xk}^2 H_z$, but also includes the term $k'(x)\partial_{kk}^2 H_z$.) Since the coefficients in (4.12) are real, the phase of Ψ is conserved and one also readily finds $|\Psi|$:

$$|\Psi|^2 = \frac{C}{V(x)} \exp\left(\int \frac{2\Gamma_z(x)}{V(x)} dx\right), \quad (4.13)$$

where the constant C is determined by the initial conditions. A straightforward generalisation to the multidimensional space \mathbf{x} gives

$$\nabla \cdot (\mathbf{V} |\Psi|^2) = 2\Gamma_z |\Psi|^2, \quad \mathbf{V} \doteq \partial_k H_z(\mathbf{x}, \mathbf{k})_{k=k(\mathbf{x})}, \quad (4.14)$$

with $\mathbf{k}(\mathbf{x}) \doteq \nabla\theta$. In particular, if \hat{D} is Hermitian, (4.12) becomes a conservation law:

$$\nabla \cdot (\mathbf{V} |\Psi|^2) = 0. \quad (4.15)$$

For stationary waves, this can be understood as energy conservation. More generally, when \mathbf{x} includes time, (4.15) represents conservation of the wave action, which is conserved even when the energy is not. For details, see, for example, Dodin (2022, § 7). A generalisation to vector waves is discussed by Dodin *et al.* (2019) and, in the broader context of MGO, in § 7.

Although often convenient, this model significantly relies on the smallness of the GO parameter (4.6). Near cutoffs, where the local λ is large (infinite) and the group velocity turns to zero, (4.13) predicts $|\Psi|^2 \sim 1/V(x) \rightarrow \infty$, which indicates limitations of said approximation. In this case, it is useful to promote (4.1) to an abstract vector equation,

$$\hat{D} |\psi\rangle = 0, \quad (4.16)$$

and seek an alternative representation of (4.16) in which the corresponding wavelength changes with the new coordinate slowly or not at all. We will discuss this in § 5, after we have introduced some more analytical tools in the next sections.

4.2. Geometrical optics of M -waves: special case

It is instructive to apply the previous approach to M -waves in particular, which can be studied using GO methods like any other waves. In this section, we consider an important special case where the new variables (q, p) are linked to H_z as follows. (Related transformations will be considered in §§ 5.3 and 5.4.)

Like in § 4.1, suppose that the reference-ray trajectory is given by some $k = k(x)$. Then, near the reference ray, one has²⁴

$$H_z(\mathbf{z}) \approx V(x)(k - k(x)), \quad (4.17)$$

²⁴Basically, this is Hayes's representation (Hayes 1973; Ruiz & Dodin 2017b) of H_z .

where $V(x) \doteq (\partial_k H_z)_{k=k(x)}$ can be interpreted as the group velocity on the reference ray. Let us choose the new variables (q, p) such that $p = H_z(x, k)$, whence k can be readily expressed as a function of (x, p) :

$$k(x, p) = k(x) + p/V(x). \quad (4.18)$$

To find the corresponding canonical coordinate q , we look for a generating function in the type-2 form, $F = F(x, p)$, such that (Goldstein, Poole & Safko 2011)

$$k(x, p) = \partial_x F(x, p), \quad q(x, p) = \partial_p F(x, p). \quad (4.19)$$

The former gives

$$F(x, p) = \int_0^x k(\tilde{x}, p) d\tilde{x}, \quad (4.20)$$

where we set the integration constant to zero by choice. (The tilde is used to distinguish the dummy integration variable \tilde{x} from the actual canonical coordinate x .) Then,

$$q(x, p) = \int_0^x \frac{\partial k(\tilde{x}, p)}{\partial p} d\tilde{x} = \int_0^x \frac{d\tilde{x}}{V(\tilde{x})} \equiv q(x). \quad (4.21)$$

Now, we can derive the corresponding approximation for M -waves. One way to do this is to perform an eikonal MT on \bar{M} and directly apply the results of §4.1. However, since we have already introduced the expansion (4.17), one might as well notice that the Weyl transform of (4.17) readily yields the following approximation of $\hat{H}_{\tilde{x}}$:

$$\hat{H}_{\tilde{x}} \approx -iV(x)\partial_x - \frac{i}{2} V'(x) - V(x)k(x). \quad (4.22)$$

Then, (3.23) leads to the following equations for $\bar{M} \equiv \bar{M}(x, q)$:

$$i\partial_q \bar{M} = \left(-iV(x)\partial_x - \frac{i}{2} V'(x) - V(x)k(x) \right) \bar{M}. \quad (4.23)$$

Also, (3.22) gives $(q - q(x))\bar{M}(x, q) = 0$, whence $\bar{M}(x, q) \propto \delta(q - q(x))$. Thus,

$$\bar{M}(x, q) = \frac{e^{i\vartheta(x)}}{\sqrt{V(x)}} \delta(q - q(x)), \quad \vartheta(x) \doteq \int_0^x k(\tilde{x}) d\tilde{x}, \quad (4.24)$$

where the normalisation is chosen such that (3.24) is satisfied:

$$\begin{aligned} \int dx \bar{M}^*(x, q') \bar{M}(x, q'') &= \int \frac{dx}{|V(x)|} \delta(q' - q(x)) \delta(q'' - q(x)) \\ &= \delta(q' - q'') \int \frac{dx}{|V(x)|} \delta(q' - q(x)) \\ &= \delta(q' - q'') \int dq \delta(q' - q) \\ &= \delta(q' - q''). \end{aligned} \quad (4.25)$$

Note that (4.24) readily provides a general GO solution to (4.1). Indeed, according to (4.16), a wave $|\psi\rangle$ is an eigenstate of $\hat{H} = \hat{p}$ corresponding to the eigenvalue $p = 0$. Then, the q -representation of this field is

$$\psi_{\hat{q}}(q) = a \langle \mathbf{e}_{\hat{q}}(q) | \mathbf{e}_{\hat{p}}(0) \rangle = \frac{a}{\sqrt{2\pi}} = \text{const.}, \quad (4.26)$$

where a is a constant amplitude and we have used (2.17). Accordingly,

$$\psi_{\hat{x}}(x) = \int dq \bar{M}(x, q) \psi_{\hat{q}}(q) = \psi_{\hat{q}} \int dq \bar{M}(x, q) = \psi_{\hat{q}} \frac{e^{i\vartheta(x)}}{\sqrt{V(x)}}. \quad (4.27)$$

This coincides with the well-known WKB solution for eikonal waves, with $\psi_{\hat{q}}$ serving as a constant factor determined by the initial conditions.

Of course, the same result can as well be obtained by directly applying GO to $\psi_{\hat{x}}$. However, the advantage of using M -waves is that doing so separates the problem of calculating the evolution of $|\psi\rangle$ from calculating its representation. Errors that may emerge from the inaccuracy of the mapping $\psi_{\hat{q}} \mapsto \psi_{\hat{x}}$ do not accumulate (except, possibly, in the phase). For example, even though (4.27) fails when V turns to zero, its validity gets reinstated as soon as V becomes large enough. In other words, one does not need to integrate the amplitude equation through singularity. One needs to solve continuously only for $\psi_{\hat{q}}$ (which is constant here but can be more complicated in general) while mapping it to $\psi_{\hat{x}}$ only occasionally, when that is easy to do with a desired accuracy.

4.3. Geometrical optics of M -waves: generic case

Let us now consider a generic case when GO M -waves are not delta-shaped, but quasimonochromatic in both x and q .²⁵ Then, one can search for a solution of (3.26) in the eikonal form, as usual:

$$\bar{M}(x, q) \equiv M^*(q, x) = e^{i\Theta(x, q)} \mathfrak{M}(x, q), \quad (4.28)$$

where Θ is a real phase (eikonal) and \mathfrak{M} is a real amplitude. Such M -waves are well studied in quantum mechanics; see, for example, Miller (1974, Section II). Here, we offer a concise derivation of those known results from an alternative perspective.

Assuming that Θ is fast compared with \mathfrak{M} , it satisfies the two ‘local dispersion relations’, or Hamilton–Jacobi equations, that flow from (3.26):

$$K(q, -\partial_q \Theta) = \partial_x \Theta, \quad (4.29a)$$

$$P(x, \partial_x \Theta) = -\partial_q \Theta. \quad (4.29b)$$

Since $K(q, -\partial_q \Theta) = K(q, p) = k$ and $P(x, \partial_x \Theta) = P(x, k) = p$, (4.29) can be expressed as

$$k = \partial_x \Theta(x, q), \quad p = -\partial_q \Theta(x, q). \quad (4.30)$$

As a function that satisfies (4.30), $\Theta(x, q)$ can be recognised as the type-1 generating function of the canonical transformation (3.4). Generating functions of other types emerge if one re-attributes the old or (and) new coordinates as the momenta, and

²⁵The relation between this case and the case considered in §4.2 is explained at the end of this section.

vice versa, with the appropriate sign changes to preserve symplecticity (§ 3.5.3). For example, in canonical variables $(q', p') \doteq (p, -q)$, (4.30) becomes

$$k = \partial_x \Theta(x, p'), \quad q' = \partial_{p'} \Theta(x, p'), \quad (4.31)$$

which makes Θ a type-2 generating function. Likewise, using $(x', k') \doteq (k, -x)$, one has

$$x' = -\partial_{k'} \Theta(k', q), \quad p = -\partial_q \Theta(k', q), \quad (4.32)$$

so Θ becomes a type-3 generating function. Finally, using both (q', p') and (x', k') , one can write

$$x' = -\partial_{k'} \Theta(k', p'), \quad q' = \partial_{p'} \Theta(k', p'), \quad (4.33)$$

in which case, Θ serves as a type-4 generating function. A reader interested in brushing up on this topic is referred to Goldstein *et al.* (2011).

Since $K(q, -i\partial_q)$ serves as the Hamiltonian for the dynamics in ‘time’ x , the symbol $K(q, p)$ serves as the ray Hamiltonian. Since this Hamiltonian does not depend on x explicitly, it is conserved on rays, $d_x K = 0$, where d_x is the convective derivative associated with the ‘group velocity’ $K_p(x, q) \doteq (\partial_p K)(q, p(x, q))$ as a field on the ‘timespace’ (x, q) . Similarly, $P(x, -i\partial_x)$ serves as the Hamiltonian for the dynamics in ‘time’ q . Since this Hamiltonian does not depend on q explicitly, it is conserved on rays, $d_q P = 0$, where d_q is now the convective derivative associated with the ‘group velocity’ $P_k(x, q) \doteq (\partial_k P)(x, k(x, q))$ as a field on the ‘timespace’ (q, x) . From

$$0 = d_x K = \partial_x k(x, q) + K_p \partial_q k(x, q), \quad (4.34a)$$

$$0 = d_q P = \partial_q p(x, q) + P_k \partial_x p(x, q), \quad (4.34b)$$

one obtains the following equations that will be useful in the next section:

$$K_p(x, q) = -\frac{\partial_x k(x, q)}{\partial_q k(x, q)} = -\frac{\Theta_{xx}}{\Theta_{xq}}, \quad (4.35a)$$

$$P_k(x, q) = -\frac{\partial_q p(x, q)}{\partial_x p(x, q)} = -\frac{\Theta_{qq}}{\Theta_{xq}}. \quad (4.35b)$$

The lower indices in $\Theta \equiv \Theta(x, q)$ denote the corresponding partial derivatives and the argument (x, q) will now be occasionally omitted for brevity.

To calculate the envelope \mathfrak{M} , let us apply the action conservation theorem (4.15) to M -waves. For the dynamics on timespace $\mathbf{x} = (x, q)$, one has $\mathbf{V} = (1, K_p)$ and $|\Psi|^2 = \mathfrak{M}^2$, so (4.15) leads to $\partial_x \mathfrak{M}^2 + \partial_q (K_p \mathfrak{M}^2) = 0$. Similarly, for the dynamics on timespace $\mathbf{x} = (q, x)$, one has $\mathbf{V} = (1, P_k)$ and $|\Psi|^2 = \mathfrak{M}^2$, so (4.15) leads to $\partial_q \mathfrak{M}^2 + \partial_x (P_k \mathfrak{M}^2) = 0$. Let us rewrite these equations as follows:

$$\partial_x \mathfrak{M}^2 + (\partial_q K_p) \mathfrak{M}^2 + K_p \partial_q \mathfrak{M}^2 = 0, \quad (4.36a)$$

$$\partial_q \mathfrak{M}^2 + (\partial_x P_k) \mathfrak{M}^2 + P_k \partial_x \mathfrak{M}^2 = 0. \quad (4.36b)$$

Substituting $\partial_q \mathfrak{M}^2$ from the second equation into the first one yields

$$\ln \mathfrak{M}^2 = \int dx \frac{K_p (\partial_x P_k) - \partial_q K_p}{1 - P_k K_p} = \ln |\Theta_{xq}| + \text{const.}, \quad (4.37)$$

where the latter equality is obtained by direct calculation using (4.35). The integration constant is independent of x . Since one can equally arrive at (4.37) by excluding $\partial_x \mathfrak{M}^2$ instead of $\partial_q \mathfrak{M}^2$ and integrating over q instead of x , this constant cannot depend on q either. Thus, the GO solution for M -waves is

$$\bar{M}(x, q) = C \sqrt{\Theta_{xq}} e^{i\Theta}, \quad (4.38)$$

where C is a constant. This factor is determined by the normalisation condition (3.25):

$$1 \approx |C|^2 \int dx ds |\Theta_{xq}(x, q)| e^{-i\Theta_q(x, q)s} = |C|^2 \int d\xi ds e^{-i\xi s} = 2\pi |C|^2, \quad (4.39)$$

so $|C| = (2\pi)^{-1/2}$. (Here, we introduced a variable transformation $x \mapsto \xi \doteq \Theta_q(x, q)$ at fixed q and used that $d\xi = \Theta_{qx} dx$.) Then, up to a piecewise-constant phase, one has

$$\bar{M}(x, q) = \sqrt{\frac{\Theta_{xq}}{2\pi}} e^{i\Theta}, \quad (4.40)$$

and, as a reminder, $\Theta_{xq} = \partial_q k = -\partial_x p$.²⁶ The initial phase is determined by the preferred gauge (§ 3.2). Once it is set, the remaining phase is determined by the branch of the square root and changes by $\pm\pi/2$ when Θ_{xq} goes through zero. This is the standard Maslov-index problem, which we will not revisit here, but see Miller (1974, pp. 85–86).

An example illustrating (4.40) will be discussed in Appendix B.2, and this formula is also useful for deriving the quantisation condition for closed orbits (Appendix D). However, keep in mind that (4.40) is constructed only as the leading-order approximation and we will also need alternative approximations of M -waves for our purposes (§ 5.4).

Finally, one might wonder why (4.40) is so different from (4.24), which seems to be based on similar approximations. The explanation is as follows. In § 4.2, we used a type-2 generating function $F(x, p)$ to obtain $q = q(x, p)$ and $k = k(x, p)$. Normally, one can invert the former to obtain $p = p(x, q)$ and $k = k(x, p(x, q))$; then, one can reformulate this as a type-1 transformation and the results of the present section apply. However, the specific transformation in § 4.2 is singular in that, by (4.21), our specific $q(x, p)$ happens to be independent of p . This makes the inversion impossible and leads to a delta-shaped M -wave (4.24).

5. Natural coordinates

Now, let us discuss how to use MTs for reduced wave modelling beyond the realm of conventional GO. This section is focused on introducing the key concepts. Specifics, such as envelope equations and examples, will be presented in § 6.

5.1. Preliminaries

5.1.1. Tangent space

Our goal here is to construct a phase-space coordinate system aligned with the reference ray. We will assume that the ray is governed by the Hamiltonian H_z and thus

²⁶For a transformation $(x, k) \mapsto (q, p)$ that represents the evolution of a classical Hamiltonian system (see the footnote 15), (4.40) is equivalent to the Van Vleck approximation for the propagator of the corresponding Schrödinger equation; for example, see Blair (2022).

satisfies $H_z(\mathbf{z}) = 0$, like in § 4.1. (However, other Hamiltonians can be advantageous or even necessary in some cases, as will be discussed in § 7.) Consider a narrow region of a fixed phase-space point \mathbf{z}_0 that satisfies $H_z(\mathbf{z}_0) = 0$. Assuming that H_z is sufficiently smooth (the quantitative condition for this will be discussed in § 5.1.2), let us consider its Taylor expansion in $\mathbf{z} - \mathbf{z}_0$:

$$H_z(\mathbf{z}) = (\mathbf{z} - \mathbf{z}_0) \cdot \underbrace{\partial_z H_z(\mathbf{z}_0)}_{-\mathbf{J}\mathbf{v}} + \mathcal{Y}, \quad \mathbf{v} \doteq \mathbf{J}\partial_z H_z(\mathbf{z}_0), \quad (5.1)$$

where we used $\mathbf{J}\mathbf{J} = -\mathbf{I}$. Here, \mathbf{v} is understood as the phase-space velocity at \mathbf{z}_0 , so we will also write it as $\mathbf{v} = \dot{\mathbf{z}}_0$. (The index 0 denotes that the corresponding quantity is evaluated at \mathbf{z}_0 .) The term \mathcal{Y} subsumes the second- and higher-order terms of the Taylor series. Albeit non-negligible, it is small at small $\mathbf{z} - \mathbf{z}_0$, so GO applicability is determined entirely by the first term, $-(\mathbf{z} - \mathbf{z}_0) \cdot \mathbf{J}\mathbf{v}$. Because this term is linear in \mathbf{z} , a linear variable transformation is generally enough to reinstate GO at least locally. This is done as follows.

Assuming the same notation as earlier, let us adopt the operator transformation (3.1) in the form

$$\hat{\mathbf{y}} = \mathbf{S}(\hat{\mathbf{z}} - \mathbf{z}_0), \quad (5.2)$$

where \mathbf{S} is a symplectic matrix. As discussed in § 3.5.3, this corresponds to the phase-space transformation $\mathbf{y}(\mathbf{z}) \equiv \mathbf{y}_z = \mathbf{S}(\mathbf{z} - \mathbf{z}_0)$ or, equivalently, $\mathbf{z}(\mathbf{y}) \equiv \mathbf{z}_y = \mathbf{z}_0 + \bar{\mathbf{S}}\mathbf{y}$, where $\bar{\mathbf{S}}$ is a symplectic matrix that is the inverse of \mathbf{S} . Let us represent $\bar{\mathbf{S}}$ through its columns \mathbf{s}_q and \mathbf{s}_p ,

$$\bar{\mathbf{S}} = (\mathbf{s}_q \quad \mathbf{s}_p), \quad (5.3)$$

and assume the notation $\mathbf{y} = (q, p)^\top$, as before. Then, q and p can be understood as the components of $\mathbf{z} - \mathbf{z}_0$ in the basis formed by the vectors \mathbf{s}_q and \mathbf{s}_p :

$$\mathbf{z} - \mathbf{z}_0 = \mathbf{s}_q q + \mathbf{s}_p p. \quad (5.4)$$

Also, the transformation (5.2) can be expressed through (5.3) using symplecticity of \mathbf{S} ; that is, inverting $\mathbf{S}\mathbf{J}\mathbf{S}^\top = \mathbf{J}$ yields $\mathbf{S} = -\mathbf{J}\bar{\mathbf{S}}^\top\mathbf{J}$ (see also (3.53)).

Let us choose \mathbf{s}_q such that it is parallel to \mathbf{v} , i.e. the q -axis is tangent to the ray,

$$\mathbf{s}_q = \mathbf{v}/v. \quad (5.5)$$

Because of this, we will call \mathbf{y} a tangent phase space. The non-zero scalar v can be anything for now (but this freedom will be removed in § 5.1.2 for nonlinear coordinate transformations, where we also comment on the physical meaning of v). It may be tempting to choose \mathbf{s}_p such that it is perpendicular to \mathbf{s}_q ; however, angles are not well defined on a symplectic space for the lack of a metric.²⁷ Instead, the only natural constraint on \mathbf{s}_p is that $\bar{\mathbf{S}}$ must be symplectic, i.e. $\mathbf{J} = \bar{\mathbf{S}}^\top\mathbf{J}\bar{\mathbf{S}}$:

$$\mathbf{J} = \begin{pmatrix} \mathbf{s}_q^\top \\ \mathbf{s}_p^\top \end{pmatrix} \mathbf{J} (\mathbf{s}_q \quad \mathbf{s}_p) = \begin{pmatrix} \mathbf{s}_q \cdot \mathbf{J}\mathbf{s}_q & \mathbf{s}_q \cdot \mathbf{J}\mathbf{s}_p \\ \mathbf{s}_p \cdot \mathbf{J}\mathbf{s}_q & \mathbf{s}_p \cdot \mathbf{J}\mathbf{s}_p \end{pmatrix} = (\mathbf{s}_q \wedge \mathbf{s}_p)\mathbf{J}. \quad (5.6)$$

²⁷Since x and k have different units, angles on the \mathbf{z} plane are unit-dependent and the linear term in (5.1) has no natural units. Natural units will be introduced in § 5.1.2 by studying \mathcal{Y} .

Hence, the requirement that $\bar{\mathbf{S}}$ must be symplectic is equivalent to the requirement that $\mathbf{s}_q \wedge \mathbf{s}_p = 1$, whence

$$\mathbf{s}_p = \mathbf{u}/u, \quad u \doteq \mathbf{v} \wedge \mathbf{u}/v. \quad (5.7)$$

We choose \mathbf{u} such that u is non-zero. Assuming the notation $\mathbf{v} = (a_1, a_2)^\top$ and $\mathbf{u} = (b_1, b_2)^\top$, one has $\mathbf{v} \wedge \mathbf{u} = a_1 b_2 - a_2 b_1$, so having non-zero u is equivalent to having \mathbf{u} linearly independent from \mathbf{v} , i.e. having (5.3) non-degenerate. Other than that, \mathbf{u} can be chosen arbitrarily (but see also § 5.1.2).

Because the transformation (5.3) is linear, it preserves the symbol of \hat{H} , so $H_y(\mathbf{y}) = H_z(\mathbf{z}) = \mathcal{Y} - (\bar{\mathbf{S}}\mathbf{y})^\top \mathbf{J}\mathbf{v}$. Notice that

$$-\mathbf{y}^\top \begin{pmatrix} \mathbf{s}_q^\top \\ \mathbf{s}_p^\top \end{pmatrix} \mathbf{J}\mathbf{v} = -\mathbf{y}^\top \begin{pmatrix} \mathbf{v} \wedge \mathbf{v}/v \\ \mathbf{u} \wedge \mathbf{v}/u \end{pmatrix} = (q \quad p) \begin{pmatrix} 0 \\ v \end{pmatrix} = vp, \quad (5.8)$$

so, in summary, $H_y(\mathbf{y}) = \mathcal{Y} + vp$. The corresponding ray equations are

$$\dot{\mathbf{y}} = \mathbf{J}\partial_y H_y \equiv \{\mathbf{y}, H_y\}, \quad (5.9)$$

or, explicitly,

$$\dot{q} = v + \partial_p \mathcal{Y}, \quad \dot{p} = -\partial_q \mathcal{Y}. \quad (5.10)$$

Since $\mathcal{Y} = \mathcal{O}((z - z_0)^2)$, these lead to $\dot{q}_0 = v$ and $\dot{p}_0 = 0$. In this sense, v can be understood as the ‘symplectic speed’ at \mathbf{z}_0 . However, v must not be confused with the length of the phase-space-velocity vector \mathbf{v} , because the vector length is generally undefined on a symplectic space (see the footnote 10).

Also note that, since \mathcal{Y} is small, the wavelength $2\pi/p$ evolves slowly. This justifies the GO approximation and (5.9) *a posteriori*. As one can also see easily, (5.9) are equivalent to the ray equations in the x -representation, (4.11). In other words, the ray equations (4.11) remain applicable even near cutoffs, where GO is inapplicable in the x -representation. In practice, this fact is often taken for granted, but it is non-trivial and requires justification. As seen from the previous argument, such ‘unreasonable effectiveness’ of the ray equations is due to the existence of a symplectic transformation $\mathbf{z} \mapsto \mathbf{y}$ that reinstates GO locally for any smooth H_z while conserving the Poisson bracket and the dispersion-operator symbol.

With GO reinstated in the q -representation, one can also search for the wavefield in the eikonal form $\psi_{\hat{q}} = e^{i\theta(q)}\Psi(q)$ and derive the envelope operator $\hat{\mathcal{H}}_y$ for the envelope Ψ in the q -space the same way the envelope equation was derived in § 4.1 in the x -space. If \mathcal{Y} is small globally, i.e. if H_z is mostly linear in \mathbf{z} (locally) at all \mathbf{z} , this envelope equation will remain sufficient indefinitely. However, such dispersion operators are rarely found in practice. A typical H_z is a nonlinear function of \mathbf{z} , so the choice of the q -representation that is convenient at some \mathbf{z} is likely to become inconvenient at other \mathbf{z} . Then, a single linear transformation is not enough.

There are two ways to deal with this problem. One is to define different linear transformations at different points $\mathbf{z}_0^{(i)}$ along the ray and use MTs to map the wavefield from the q -space at $\mathbf{z}_0^{(i)}$ to the q -space at $\mathbf{z}_0^{(i+1)}$ (figure 1a). For example, one can choose $\mathbf{z}_0^{(i+1)}$ to be infinitesimally close to $\mathbf{z}_0^{(i)}$; then, the MT for remapping the field will be a near-identity transformation, which can be convenient. This approach is summarised in (Lopez & Dodin 2022). The alternative is to use just one but

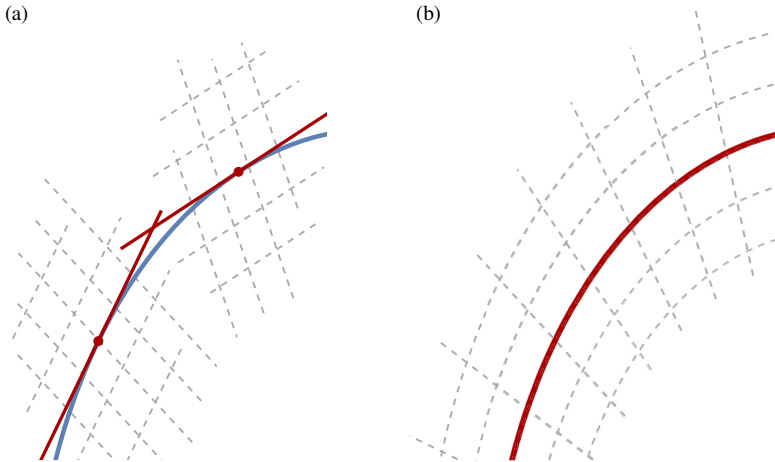


FIGURE 1. Two ways of constructing the y -coordinate grids around a reference ray (blue; coincides with the dispersion surface) in the z -space. (a) Grids are constructed in patches near predefined locations (red dots) on the ray via linear variable transformations. The q -axes (red) are tangent to the ray. (b) A single grid is constructed via a nonlinear variable transformation. The coordinate axis (red) coincides with the ray indefinitely.

nonlinear transformation such that the new coordinate space is aligned with the ray indefinitely (figure 1b). No field remapping is needed in this case other than to initialise the field in the new coordinate space and to output the final results back to the x -space. This is the approach that we discuss below.

5.1.2. Ray path as a coordinate

Let us consider a generic ray that is curved, i.e. has $\ddot{\mathbf{z}}$ linearly independent from $\dot{\mathbf{z}}$ in the region of interest. (Straight rays can be handled as a limit; see later.) Then, at any given \mathbf{z}_0 on the ray, the vectors

$$\mathbf{v} = \dot{\mathbf{z}}_0, \quad \mathbf{u} = \ddot{\mathbf{z}}_0 \quad (5.11)$$

form a symplectic basis²⁸

$$\mathbf{s}_q = \mathbf{v}/v, \quad \mathbf{s}_p = \mathbf{u}/u, \quad \mathbf{s}_q \wedge \mathbf{s}_p = 1, \quad (5.12)$$

where $u \doteq \mathbf{v} \wedge \mathbf{u}/v$ and the scalar v remains to be defined. This basis can be used to construct new ray-aligned coordinates $r = (\tau, h)$ (here, ‘ r ’ stands for ‘ray’) such that the ray itself would be an isoline of the new momentum h . Then, one can derive the envelope equation as a partial differential equation (PDE) in the new-coordinate representation. The envelope remains slow in this representation because the reference wavevector satisfies the dispersion relation indefinitely and there are no cutoffs in this representation by construction of r .

Since we are interested only in the local dynamics on the ray, it will be sufficient for us to find $r(\mathbf{z})$ and said PDE near a predefined generic \mathbf{z}_0 on the ray.

²⁸A similar approach is used by Kamran, Olver & Tenenblat (2009). Also note that here we assume that \mathbf{z} is two-dimensional for simplicity. In the case of a higher-dimensional phase space \mathbf{Z} , one can think of our z -space as the osculating plane spanned by $\dot{\mathbf{z}}_0$ and $\ddot{\mathbf{z}}_0$ (Tracy *et al.* 2014, Section 6.5.1).

Then, it is convenient to search for the mapping $\mathbf{z} \mapsto \mathbf{r}$ as a superposition of two transformations:

$$\underbrace{(x, k)^\top}_{\mathbf{z}} \mapsto \underbrace{(q, p)^\top}_{\mathbf{y}} \mapsto \underbrace{(\tau, h)^\top}_{\mathbf{r}}. \quad (5.13)$$

The first one, $\mathbf{y} = \bar{\mathbf{S}}(\mathbf{z} - \mathbf{z}_0)$, is the same as in § 5.1.1, except now \mathbf{u} is specified by (5.11); i.e. the q -axis is tangent to the ray at $\mathbf{z} = \mathbf{z}_0$. The second transformation, $\mathbf{y} \mapsto \mathbf{r}$, is nonlinear and modifies the coordinate grid such that the new coordinate axis coincides with the ray. (The momentum axis is also modified accordingly, such that the canonical commutation relation (3.2) is preserved.) The advantage of the splitting (5.13) is that the mapping $\mathbf{z} \mapsto \mathbf{y}$ is easy to construct exactly and, for smooth rays, $\mathbf{y} \mapsto \mathbf{r}$ can also be found analytically, at least asymptotically. In the following, we show how to do this step by step.

5.2. ‘Osculating’ harmonic oscillator

Let us start with a simplified model, where H_z can be adequately approximated (locally) with its second-order Taylor expansion in $\mathbf{z} - \mathbf{z}_0$:²⁹

$$H_z(\mathbf{z}) = -(\mathbf{z} - \mathbf{z}_0) \cdot \mathbf{J}\mathbf{v} + \frac{1}{2} (\mathbf{z} - \mathbf{z}_0) \cdot \mathbf{g}(\mathbf{z} - \mathbf{z}_0). \quad (5.14)$$

Here, $\mathbf{v} \doteq \mathbf{J}\partial_z H_z(\mathbf{z}_0) \equiv \dot{\mathbf{z}}_0$, as earlier, and \mathbf{g} is a symmetric matrix given by $\mathbf{g} \doteq (\partial_z^2 H)_0$. A related description of this model can also be found in Tracy *et al.* (2014, § 5.3.3) for a specific form of H_z . Here, we allow for a general H_z to keep the theory symplectically invariant, at least in its final form discussed in § 5.4.

Equation (5.14) leads to the following Hamilton’s equations:

$$\dot{\mathbf{z}} = \mathbf{v} + \mathbf{J}\mathbf{g}(\mathbf{z} - \mathbf{z}_0), \quad \ddot{\mathbf{z}} = \mathbf{J}\mathbf{g}\dot{\mathbf{z}}, \quad (5.15)$$

so $\ddot{\mathbf{z}}_0 = (\mathbf{J}\mathbf{g}\dot{\mathbf{z}})_0 = \mathbf{J}\mathbf{g}\mathbf{v}$. Then, $\mathbf{u} = \mathbf{J}\mathbf{g}\mathbf{v}$ and $u = -\varepsilon/v$. Here, v is a normalisation factor that was introduced in (5.5) and remains to be specified. Also,

$$\varepsilon \doteq (\ddot{\mathbf{z}} \wedge \dot{\mathbf{z}})_0 = (\mathbf{J}\mathbf{g}\mathbf{v})^\top \mathbf{J}\mathbf{v} = \mathbf{v} \cdot \mathbf{g}\mathbf{v}, \quad (5.16)$$

where we used $\mathbf{J}^\top \mathbf{J} = \mathbf{I}$ and $\mathbf{g}^\top = \mathbf{g}$. Also note that

$$\mathbf{v} \cdot \mathbf{g}\mathbf{u} = \mathbf{v} \cdot \mathbf{g}\mathbf{J}\mathbf{g}\mathbf{v} = (\mathbf{g}\mathbf{v}) \wedge (\mathbf{g}\mathbf{v}) = 0, \quad (5.17)$$

$$\mathbf{u} \cdot \mathbf{g}\mathbf{u} = (\mathbf{J}\mathbf{g}\mathbf{v})^\top \mathbf{g}(\mathbf{J}\mathbf{g}\mathbf{v}) = \mathbf{v}^\top \mathbf{g}^\top \mathbf{J}^\top \mathbf{g}\mathbf{J}\mathbf{g}\mathbf{v} = \mathbf{v} \cdot \mathbf{g}(-\mathbf{J}\mathbf{g}\mathbf{J}\mathbf{g})\mathbf{v} = g\varepsilon, \quad (5.18)$$

where we used $\mathbf{J}^\top = -\mathbf{J}$, and, for 2×2 matrices, $-\mathbf{J}\mathbf{g}\mathbf{J}\mathbf{g} = g\mathbf{I}$, where $g \doteq \det \mathbf{g}$. Similarly, the time derivative of $\ddot{\mathbf{z}} \wedge \dot{\mathbf{z}}$ is

$$\ddot{\mathbf{z}} \wedge \dot{\mathbf{z}} = (\mathbf{J}\mathbf{g}\ddot{\mathbf{z}}) \wedge \dot{\mathbf{z}} = \ddot{\mathbf{z}}^\top \mathbf{g}^\top \mathbf{J}^\top \mathbf{J}\dot{\mathbf{z}} = \ddot{\mathbf{z}}^\top \mathbf{g}\dot{\mathbf{z}} = \ddot{\mathbf{z}}^\top \mathbf{J}^\top (\mathbf{J}\mathbf{g}\dot{\mathbf{z}}) = -\ddot{\mathbf{z}} \wedge \dot{\mathbf{z}} = 0, \quad (5.19)$$

so $\ddot{\mathbf{z}} \wedge \dot{\mathbf{z}} = (\ddot{\mathbf{z}} \wedge \dot{\mathbf{z}})_0$. This means that, to the extent that the model (5.14) is applicable, one can replace the definition (5.16) with $\varepsilon = \ddot{\mathbf{z}} \wedge \dot{\mathbf{z}}$.

²⁹Limitations of this approximation are discussed at the end of this section. Also, alternative expansions will be introduced in §§ 5.3.1 and 5.4.

Let us use these results to see how (5.14) is transformed by the mapping $\mathbf{z} \mapsto \mathbf{y}$. The first term in (5.14) becomes $p v$, like before. The second term becomes

$$\begin{aligned} \frac{1}{2} (\mathbf{z} - \mathbf{z}_0) \cdot \mathbf{g}(\mathbf{z} - \mathbf{z}_0) &= \frac{1}{2} \mathbf{y}^\top \begin{pmatrix} \mathbf{v}^\top/v \\ \mathbf{u}^\top/u \end{pmatrix} \mathbf{g} \begin{pmatrix} v/v & u/u \end{pmatrix} \mathbf{y} \\ &= \frac{1}{2} \mathbf{y} \cdot \begin{pmatrix} \mathbf{v} \cdot \mathbf{g} \mathbf{v}/v^2 & \mathbf{v} \cdot \mathbf{g} \mathbf{u}/vu \\ \mathbf{u} \cdot \mathbf{g} \mathbf{v}/vu & \mathbf{u} \cdot \mathbf{g} \mathbf{u}/u^2 \end{pmatrix} \mathbf{y} \\ &= \frac{1}{2} \mathbf{y} \cdot \begin{pmatrix} \varepsilon/v^2 & 0 \\ 0 & g v^2/\varepsilon \end{pmatrix} \mathbf{y}. \end{aligned} \quad (5.20)$$

Assuming the notation $\sigma_f \doteq \text{sgn } f$ for any f , let us introduce $\Omega \doteq \sigma_\varepsilon \sqrt{|g|}$ and choose $v = (\varepsilon^2/|g|)^{1/4}$, so that the absolute values of the diagonal elements in the matrix in (5.20) are the same. Then, $\varepsilon/v^2 = |g|v^2/\varepsilon = \Omega$, so $v = |\varepsilon/\Omega|^{1/2}$, $u = -\sigma_\varepsilon |\varepsilon \Omega|^{1/2}$ and

$$\begin{pmatrix} \varepsilon/v^2 & 0 \\ 0 & g v^2/\varepsilon \end{pmatrix} = \Omega \begin{pmatrix} 1 & 0 \\ 0 & \sigma_g \end{pmatrix}. \quad (5.21)$$

Then,

$$H_y(\mathbf{y}) = H_z(\mathbf{z}) = \frac{\Omega}{2} (q^2 + \sigma_g (p + R)^2 - \sigma_g R^2), \quad (5.22)$$

where $R \doteq \sigma_g v/\Omega$, so $|R| = |v/\Omega| = |\varepsilon|^{1/2} |g|^{-3/4}$.³⁰ One can recognise this procedure as the Williamson diagonalisation of \mathbf{g} (at least, for $g > 0$), with Ω being the symplectic eigenvalue of \mathbf{g} (Nicacio 2021). Note that Ω and R are symplectic invariants by definition, and so are ε , g and v . Also, it is easily seen that R is dimensionless and the other quantities have the following units (the square brackets here stand for ‘units of’):

$$[\varepsilon] = [H_z]^3, \quad [g] = [H_z]^2, \quad [\Omega] = [H_z], \quad [v] = [H_z]. \quad (5.23)$$

Also useful for us will be the following formula:

$$|\varepsilon|^{1/3} = |\Omega| |R|^{2/3} = v/|R|^{1/3}. \quad (5.24)$$

The phase-space variables $(q, p)^\top \equiv \mathbf{y}$ can be understood as a local approximation to the natural canonical variables of \hat{H} . The matrix \mathbf{g} serves as a natural (pseudo)metric induced by \hat{H} on the symplectic space; then, $\varepsilon = \|\mathbf{v}\|_{\mathbf{g}}^2$ can be understood as the square of \mathbf{v} in this metric. The ray trajectories in the \mathbf{y} -space satisfy $H_y(\mathbf{y}) = 0$, so, within the approximation (5.22), they are conic sections: circles at $g > 0$, parabolas at $g = 0$ and hyperbolas at $g < 0$. However, we are interested only in small q and p , in which case σ_g has little effect on the ray trajectories (figure 2).³¹

³⁰Notably, $|R|^{2/3} = |\mathbf{z}'' \wedge \mathbf{z}'|^{1/3} \equiv |\mathbf{s}'|$, where the primes denote derivatives with respect to the phase Ωt and \mathbf{s} is the ‘symplectic arc length’ introduced by Kamran *et al.* (2009).

³¹Strictly speaking, at $g < 0$, the dispersion curve has two branches, potentially allowing for mode conversion, i.e. energy tunnelling from one dispersion surface to another (§7.3). However, this is largely irrelevant for us here, as our goal in this section is to construct phase-space variables around a given reference ray rather than solve a wave equation.

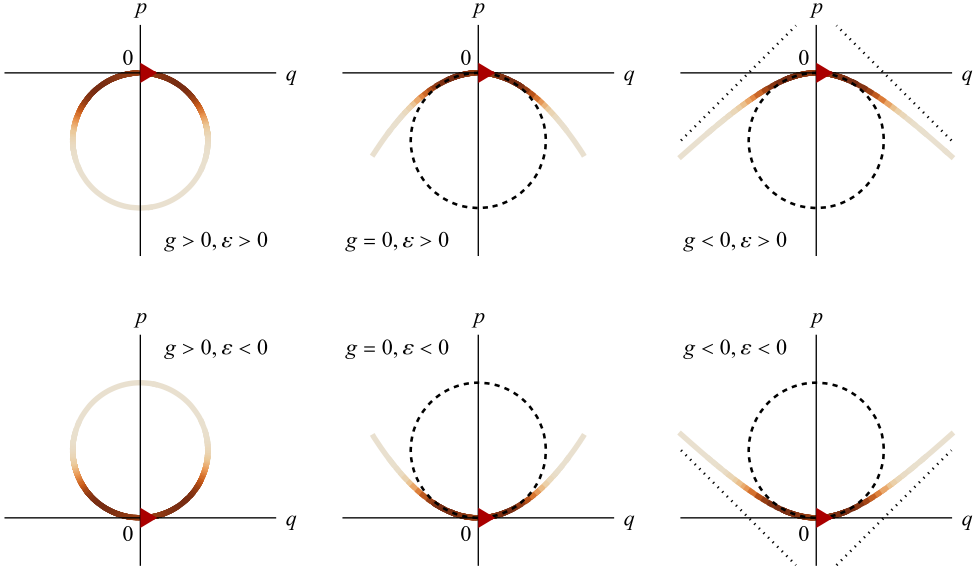


FIGURE 2. Ray trajectories $H_y(\mathbf{y}) = 0$ for various signs of g and ε . Darker colours mark, loosely, the areas of where the assumed model is adequate. Lighter colours mark the areas that are beyond the validity domain of the model. The arrows mark the direction of the ray propagation, which is always towards positive q . In the two right columns, dashed are the osculating circles. The radius of each circle is $|R| = |v/\Omega|$. Dotted are the asymptotes of the hyperbolas.

Hence, locally, the ray dynamics can be thought of as the dynamics of a harmonic oscillator with the Hamiltonian

$$H_y(\mathbf{y}) = \frac{\Omega}{2} (q^2 + (p + R)^2 - R^2). \quad (5.25)$$

We call it the ‘osculating’ harmonic oscillator (OHO). (To a dynamical system, an OHO basically is what an osculating circle is to a curve.) The phase-space orbit of OHO is a circle with radius $|R|$. Since our approximation (5.14) is generally adequate only on scales small compared with this scale, our model is applicable only when both q and p are much less than R .

The parameter Ω can be understood as OHO’s angular frequency. At $\Omega > 0$, or $\varepsilon > 0$, the phase-space trajectory is oriented clockwise, as usual. (Note that the canonical phase of the oscillator normally grows when the polar angle on the (q, p) plane decreases.) At $\Omega < 0$, or $\varepsilon < 0$, the trajectory is counter-clockwise. At $\Omega = 0$, which corresponds to $\varepsilon = 0$, the trajectory is flat, i.e. $R = \infty$. Hence, it is convenient to classify the ray trajectories in terms of their (dimensionless) ‘symplectic curvature’

$$\kappa \doteq \frac{1}{R} = \frac{\Omega}{v} = \sigma_\varepsilon \frac{|g|^{3/4}}{|\varepsilon|^{1/2}}, \quad (5.26)$$

and R will be called the symplectic radius, accordingly. (Here and further, $\text{sgn}R \equiv \text{sgn} \kappa = \sigma_\varepsilon$.) Note that the symplectic curvature can be very different from the regular curvature of a trajectory in a metric space. In particular, one can easily show that

$$\varepsilon = -\frac{d^2x}{dk^2} \left(\frac{\partial H_z}{\partial x} \right)^3 = -\frac{d^2k}{dx^2} \left(\frac{\partial H_z}{\partial k} \right)^3, \quad (5.27)$$

where the full derivatives are taken along the ray trajectory, $H_z(\mathbf{z}) = 0$. Thus, at inflection points (where $\varepsilon \rightarrow 0$), κ is infinite, whereas the regular curvature would be zero.

Finally, notice the following. The OHO approximation (5.25) for H_z corresponds to the following approximation of \hat{H} itself: $\hat{H} \approx (\Omega/2)(\hat{q}^2 + (\hat{p} + R)^2 - R^2)$. This means that, in the absence of dissipation, the wave (4.16) becomes the equation of a quantum harmonic oscillator (QHO) with energy $\Omega R^2/2$, in the natural units that correspond to unit Planck constant, $\hbar = 1$. Since the QHO energy can be expressed through its quantum mode number n as $\Omega(n + 1/2)$, one obtains

$$n = \frac{R^2}{2} - \frac{1}{2}. \quad (5.28)$$

This readily determines applicability conditions for GO in the q -representation in the region $q \ll R$ in which we are interested. Indeed, as commonly known, one can adequately approximate the dynamics of such QHO using GO at $n \gg 1$, i.e. $R \gg 1$. This makes κ a natural small parameter for MGO. One might be surprised that this parameter becomes infinite at inflection points, where rays become straight and thus, supposedly, MGO should work only better. However, notice that, near inflection points, R evolves faster than the ray moves along the OHO trajectory *per se* and, thus, the very concept of the orbit radius becomes irrelevant. In other words, the quadratic approximation (5.14) is inapplicable in such regimes and one should either include cubic terms or simply ignore the effects produced by the quadratic terms, in which case MGO is reinstated. What matters then is not the local R *per se*, but the characteristic R . We will return to this subject in §5.4.

5.3. Parabolic approximation

There is more than one way to define the transformation $\mathbf{y} \mapsto \mathbf{r}$ near a given \mathbf{z}_0 . One set of obvious variables to try are the angle–action variables of the OHO. However, as discussed in Appendix B, this approach is not optimal, so let us try a different one.

5.3.1. Operator transformation

As pointed out in §5.2, the OHO approximation is valid only at $q \ll R$, where $p \sim q^2/R \ll q$. Because of this, one can neglect p^2 compared with q^2 in (5.22) and obtain the following ‘parabolic’ approximation of the ray Hamiltonian:³²

$$H_y(\mathbf{y}) = H_z(\mathbf{z}) \approx pv + \frac{\Omega}{2} q^2. \quad (5.29)$$

(Note that this holds for any signs of g and ε , so the discontinuous function sgn never even appears in this formulation.) Let us now introduce new canonical variables $(\tau, h)^\top \equiv \mathbf{r}$ such that $h = pv + \Omega q^2/2$. To find a suitable τ , let us search for the corresponding generating function F in the type-2 form, $F = F(q, h)$ (Goldstein *et al.* 2011):

$$p(q, h) = \partial_q F(q, h), \quad \tau(q, h) = \partial_h F(q, h). \quad (5.30)$$

³²One can cast (5.29) in an even simpler, one-parameter form, $H_Y = P + \varepsilon Q^2/2$, by introducing yet another set of canonical variables, $Y \equiv (Q, P)^\top \doteq (q/v, pv)^\top$.

Since $p(q, h) = (h - \Omega q^2/2)/v$, one has $F(q, h) = (hq - \Omega q^3/6)/v$ (up to an arbitrary function of h , which we choose to be zero), so

$$\tau = q/v, \quad h = pv + \Omega q^2/2, \quad (5.31)$$

where we have repeated the expression for h for completeness. Since v and w are defined as values at a fixed \mathbf{z}_0 , the corresponding operators are given simply by

$$\hat{\tau} = \hat{q}/v, \quad \hat{h} = \hat{p}v + \Omega \hat{q}^2/2. \quad (5.32)$$

(Basically, (5.32) is just a rescaled and truncated small- p approximation of (B4).) One can readily see that they exactly satisfy the canonical commutation relation $[\hat{\tau}, \hat{h}] = i$.

5.3.2. M -waves

In addition to avoiding infinite series, the previous model is advantageous in that it allows one to explicitly calculate M -waves in a closed form. Indeed, (3.23) for $\bar{M} \doteq M_{\hat{q} \leftarrow \hat{\tau}}$ becomes an equation solvable analytically:

$$i\partial_\tau \bar{M} = -iv\partial_q \bar{M} + \frac{1}{2} wq^2 \bar{M} \quad (5.33)$$

(and (3.19) is just the complex conjugate thereof). Since $\hat{\tau} = \hat{q}/v$, one can adopt $|\mathbf{e}_\tau(\tau)\rangle = \sqrt{v} |\mathbf{e}_{\hat{q}}(v\tau)\rangle$; then, the initial condition for (5.33) is

$$\bar{M}(q, 0) \equiv \langle \mathbf{e}_{\hat{q}}(q) | \mathbf{e}_\tau(0) \rangle = \sqrt{v} \langle \mathbf{e}_{\hat{q}}(q) | \mathbf{e}_{\hat{q}}(0) \rangle = \sqrt{v} \delta(q). \quad (5.34)$$

The corresponding solution of (5.33) is

$$\bar{M}(q, \tau) = \frac{1}{\sqrt{v}} \delta\left(\tau - \frac{q}{v}\right) \exp\left(-\frac{i\Omega q^3}{6v}\right). \quad (5.35)$$

Accordingly, the pseudo-measure (3.29), with $M \equiv \bar{M}^*$, is given by

$$\begin{aligned} \mu &= \frac{1}{2\pi} \int ds' \int ds M(\tau + s/2, q + s'/2) M^*(\tau - s/2, q - s'/2) e^{-ihs + ips'} \\ &= \frac{1}{2\pi v} \int ds \int ds' \delta\left(\tau - \frac{q}{v} + \frac{s - s'/v}{2}\right) \delta\left(\tau - \frac{q}{v} - \frac{s - s'/v}{2}\right) \\ &\quad \times \exp\left(\frac{i\Omega}{6v} \left(q + \frac{s'}{2}\right)^3 - \frac{i\Omega}{6v} \left(q - \frac{s'}{2}\right)^3 - ihs + ips'\right) \\ &= \frac{1}{2\pi v} \int ds \int ds' \delta\left(\tau - \frac{q}{v}\right) \delta\left(s - \frac{s'}{v}\right) \\ &\quad \times \exp\left(\frac{i\Omega}{6v} \left(q + \frac{s'}{2}\right)^3 - \frac{i\Omega}{6v} \left(q - \frac{s'}{2}\right)^3 - ihs + ips'\right) \\ &= \frac{1}{2\pi} \int ds \delta\left(\tau - \frac{q}{v}\right) \\ &\quad \times \exp\left(\frac{i\Omega}{6v} \left(q + \frac{sv}{2}\right)^3 - \frac{i\Omega}{6v} \left(q - \frac{sv}{2}\right)^3 + is(pv - h)\right) \\ &= \frac{1}{2\pi} \int ds \delta\left(\tau - \frac{q}{v}\right) \exp\left(\frac{i\epsilon s^3}{24} + is(H_y(\mathbf{y}) - h)\right), \end{aligned} \quad (5.36)$$

where we used $\Omega v^2 = \varepsilon$. This result can also be expressed as

$$\mu(\mathbf{r}, \mathbf{y}) = \text{Ai}_\varepsilon(H_\gamma(\mathbf{y}) - h) \delta(\tau - q/v), \quad (5.37)$$

where we have introduced the following (real) function, with γ as a real parameter:

$$\text{Ai}_\gamma(z) \doteq \frac{1}{2\pi} \int_{-\infty}^{\infty} dt e^{izt + i\gamma t^3/24} = \frac{2}{|\gamma|^{1/3}} \text{Ai}\left(\frac{2}{|\gamma|^{1/3}} \sigma_\gamma z\right). \quad (5.38)$$

Here, $\sigma_\gamma \doteq \text{sgn}\gamma$, and Ai is the Airy function of the first kind. Importantly (Appendix C),

$$\lim_{\gamma \rightarrow 0} \text{Ai}_\gamma(z) = \delta(z). \quad (5.39)$$

As to be seen, the function Ai_γ , which we will call a rescaled Airy function (RAF), plays a fundamental role in MGO and emerges in multiple contexts.

5.4. General case

Let us now present a formulation that is not restricted to the quadratic approximation (5.14) and also subsumes the standard WKB approximation (4.27) as a special case.

5.4.1. Operator transformation

Let us start with assuming a generic H_z in the form

$$H_z(\mathbf{z}) = \mathcal{H}(x/\Delta x, k/\Delta k), \quad (5.40)$$

where Δx and Δk are the characteristic scales of the system in x and k , respectively, and \mathcal{H} is some function with order-one scales. Consider new dimensionless variables

$$X \doteq Rx/\Delta x, \quad K \doteq Rk/\Delta k, \quad (5.41)$$

where R is a dimensionless constant. We require that the transformation $\mathbf{z} \mapsto \mathbf{Z} \equiv (X, K)^\top$ be symplectic; then, $H_z(\mathbf{Z}) = H_z(\mathbf{z})$ and

$$R = \sqrt{\Delta x \Delta k}. \quad (5.42)$$

We will call this quantity a symplectic scale. To the extent that the quadratic approximation (5.14) is adequate at least locally, R coincides with the absolute value of the symplectic radius that was introduced in §5.2. For this reason, in the following, we occasionally mix the terms ‘symplectic scale’ and ‘symplectic radius’, and also refer to the inverse symplectic scale as symplectic curvature.

Much like in §4.1, our goal now is to derive field equations that are accurate to the first (not zeroth) order in the ‘MGO parameter’

$$\epsilon \doteq R^{-2} \equiv (\Delta x \Delta k)^{-1} \ll 1. \quad (5.43)$$

It is similar to the GO parameter introduced in §4.1, but the absolute value of k is now replaced with the spectral-inhomogeneity scale Δk ; i.e. the smallness of the

local wavelength *per se* is irrelevant. To the extent that the OHO model is applicable, ϵ coincides with κ^2 , where κ is the symplectic curvature introduced in § 5.2.

Let us perform a linear canonical transformation $\mathbf{Z} \mapsto \mathbf{y} \equiv (q, p)^\top$, as in § 5.1.1, so that the q -axis is along \mathbf{v} that is tangent to the ray. Then, $H_y(\mathbf{y}) = \mathcal{H}(q/R, p/R)$, so the ray trajectory satisfies $\mathcal{H}(q/R, p/R) = 0$. Its solution (for a given branch) can be written in the form $p = R\mathcal{P}(q/R) \equiv p(q)$, where \mathcal{P} is an order-one function that has order-one scales. (If H_z is linear in \mathbf{z} , then $\mathcal{P} = 0$ by construction. If H_z is quadratic, then $p(q) = -R + \sqrt{R^2 - q^2}$, so $p(q) \approx q^2/2R$ at $q \ll R$, as in § 5.3.) By Taylor-expanding $H_y(\mathbf{y})$ in $p - p(q)$, one obtains

$$H_y(\mathbf{y}) \approx V(q)(p - p(q)), \quad (5.44)$$

where $V(q) \doteq (\partial_p H_y)_{p=p(q)}$, so $V(0) = v$. Like (4.17), (5.44) can be understood as Hayes's representation of H_y . For GO, this representation is precise in that it leads to the correct ray equations and correctly captures the leading-order amplitude dynamics, because those are entirely determined by the first-order derivatives of the symbol. The MT that follows from (5.44) will not be exact, because $\mathcal{O}([p - p(q)]^2)$ terms have been neglected. Still, assuming $p - p(q) = \mathcal{O}(q^2)$, (5.44) differs from the exact H_y by at most $\mathcal{O}(q^4)$, which is assumed negligible (at not-too-large q that we are interested in locally).

Let us perform a canonical transformation $\mathbf{y} \mapsto \mathbf{r} \equiv (\tau, h)^\top$ such that $h = H_y$, i.e.³³

$$p(q, h) = p(q) + h/V(q). \quad (5.45)$$

Like in § 5.3, to find what the new coordinate τ should be to keep the transformation canonical, we look for a generating function F in the type-2 form:

$$p(q, h) = \partial_q F(q, h), \quad \tau(q, h) = \partial_h F(q, h). \quad (5.46)$$

The former gives

$$F(q, h) = \int_0^q p(\tilde{q}, h) d\tilde{q}, \quad (5.47)$$

where we set the integration constant to zero by choice. (The tilde is used to distinguish the dummy integration variable \tilde{q} from the actual canonical coordinate q .) Then,

$$\tau(q, h) = \int_0^q \frac{\partial p(\tilde{q}, h)}{\partial h} d\tilde{q} = \int_0^q \frac{d\tilde{q}}{V(\tilde{q})} \equiv \tau(q). \quad (5.48)$$

In other words, τ and h are the ray time and the ray energy. (By the ray energy, we mean the value of the ray Hamiltonian, not the wave energy that depends on the wave amplitude.) The corresponding operators are

$$\hat{\tau} = \tau(\hat{q}), \quad \hat{h} = \frac{1}{2}(\hat{V}\hat{p} + \hat{p}\hat{V}) - \hat{V}p(\hat{q}), \quad (5.49)$$

³³Remember that $p(q)$ is a prescribed function determined by the shape of the ray trajectory, and $p(q, h)$ is the new momentum variable p expressed as a function of q and h . The same notation, p , is justified for both in that $p(q, h)$ and p coincide on the ray; i.e. one can consider $p(q)$ as a shorthand for $p(q, 0)$.

where $\hat{V} \doteq V(\hat{q})$ (note, though, that \hat{p} is not the same as $p(\hat{q})$), so $[\hat{\tau}, \hat{V}] = 0$ and

$$\begin{aligned}
 [\hat{\tau}, \hat{h}] &= [\hat{\tau}, \hat{V}\hat{p} + \hat{p}\hat{V}]/2 \\
 &= (\hat{\tau}\hat{V}\hat{p} - \hat{V}\hat{p}\hat{\tau} + \hat{\tau}\hat{p}\hat{V} - \hat{p}\hat{V}\hat{\tau})/2 \\
 &= (\hat{\tau}\hat{V}\hat{p} - \hat{V}[\hat{p}, \hat{\tau}] - \hat{V}\hat{\tau}\hat{p} + \hat{\tau}[\hat{p}, \hat{V}] + \hat{\tau}\hat{V}\hat{p} - \hat{p}\hat{V}\hat{\tau})/2 \\
 &= ([\hat{\tau}\hat{V}, \hat{p}] - \hat{V}[\hat{p}, \hat{\tau}] + \hat{\tau}[\hat{p}, \hat{V}])/2 \\
 &= i((\tau(q)V(q))' + V(q)\tau'(q) - \tau(q)V'(q))/2 \\
 &= iV(q)\tau'(q) \\
 &= i.
 \end{aligned} \tag{5.50}$$

5.4.2. *M*-waves

The approximation (5.44) corresponds to the following approximation of the operator $\hat{H}_{\hat{q}}$ (cf. § 4.1):

$$\hat{H}_{\hat{q}} \approx -iV(q)\partial_q - \frac{i}{2} V'(q) - V(q)p(q). \tag{5.51}$$

Then, (3.22) and (3.23) lead to the following equations for $\bar{M} \equiv \bar{M}(q, \tau)$:

$$(\tau - \tau(q))\bar{M} = 0, \tag{5.52}$$

$$i\partial_\tau \bar{M} = \left(-iV(q)\partial_q - \frac{i}{2} V'(q) - V(q)p(q) \right) \bar{M}. \tag{5.53}$$

These can be readily solved:

$$\bar{M}(q, \tau) = \frac{e^{i\vartheta(q)}}{\sqrt{V(q)}} \delta(\tau - \tau(q)), \quad \vartheta(q) \doteq \int_0^q p(\tilde{q}) d\tilde{q}, \tag{5.54}$$

which generalises (5.35) to non-constant $V(q)$ and arbitrary $p(q)$. To the extent that (5.44) is valid, this result is exact. One can also notice parallels with (4.24), which is the same formula applied to different coordinates.

To calculate the generalisation of the pseudo-measure (5.37), notice that

$$\tau \left(q \pm \frac{s'}{2} \right) \approx \tau(q) \pm \frac{s'}{2V} - \frac{V'}{8V^2} s'^2, \tag{5.55}$$

where $V \equiv V(q)$ and $V'(q) \equiv dV(q)/dq$, while the prime in s' is only a notation that distinguishes said variable from s . Then,

$$\begin{aligned}
 &\delta \left(\tau + \frac{s}{2} - \tau \left(q + \frac{s'}{2} \right) \right) \delta \left(\tau - \frac{s}{2} - \tau \left(q - \frac{s'}{2} \right) \right) \\
 &\approx \delta \left(\tau - \tau(q) + \frac{s - s'/V}{2} + \frac{V's'^2}{8V^2} \right) \delta \left(\tau - \tau(q) - \frac{s - s'/V}{2} + \frac{V's'^2}{8V^2} \right) \\
 &= \delta \left(\tau - \tau(q) + \frac{V's'^2}{8V^2} \right) \delta \left(s - \frac{s'}{V} \right) \\
 &= \delta \left(\tau - \tau(q) + \frac{V's^2}{8} \right) \delta \left(s - \frac{s'}{V} \right) \\
 &\approx \left(1 + \frac{V's^2}{8} \frac{\partial}{\partial \tau} \right) \delta(\tau - \tau(q)) \delta \left(s - \frac{s'}{V} \right),
 \end{aligned} \tag{5.56}$$

where the last equality is based on the first-order Taylor expansion. (Admittedly, Taylor-expanding a delta function is a questionable procedure, but it makes sense as a shorthand for Taylor-expanding what actually matters, namely, integrals of delta functions.) Then,

$$\mu = \left[1 - (V'/8) \partial_{\tau hh}^3\right] \mu_0, \quad \mu_0 = \mathfrak{A} \delta(\tau(q) - \tau), \quad (5.57)$$

where \mathfrak{A} is a generalisation of (5.38):

$$\mathfrak{A} \doteq \frac{1}{2\pi} \int ds e^{i\phi}, \quad \Phi \doteq (pV - h)s - \int_{q-sV/2}^{q+sV/2} p(\tilde{q}) d\tilde{q}. \quad (5.58)$$

Remember that p is an independent variable, $V \equiv V(q)$ and $p(q)$ is a function introduced in § 5.4.1.

5.4.3. Symplectically invariant model for μ

Since we are already neglecting terms of order V'' , the factor V' in (5.57) can be considered commuting with the derivatives. Then, using (3.41) and integrating by parts, one obtains the following formula for remapping symbols between the y - and r -representations:

$$O_y(y) = \int \left(O_r(r) + \frac{V'}{8} \partial_{\tau hh}^3 O_r(r) \right) \mu_0 d\tau dh. \quad (5.59)$$

In the following, we will be interested in O_r that are either τ -independent or smooth, namely, have symplectic scales of order R or larger. In either case, the second term in (5.59) is $\mathcal{O}(R^{-4})$, i.e. $\mathcal{O}(\epsilon^2)$, which is considered negligible. Thus, in the following, we adopt $\mu \approx \mu_0$, i.e.

$$\mu \approx \mathfrak{A} \delta(\tau(q) - \tau). \quad (5.60)$$

One can also show (Appendix C) that, at $R \gg 1$, one can approximate \mathfrak{A} as

$$\mathfrak{A} \approx \text{Ai}_\epsilon(H_y(y) - h) \quad (5.61)$$

with $\epsilon = -V^3 p''$. In fact, even when (5.43) is not satisfied, yet all higher-order derivatives of $p(q)$ are small enough, \mathfrak{A} can still be approximated with a delta function, $\mathfrak{A} \approx \delta(H_y(y) - h)$, which may be sufficiently accurate for practical purposes. In this case, \mathfrak{A} can be safely replaced with Ai_ϵ because the latter can be approximated with a delta function as well. Hence, we will assume (5.61) either literally, when (5.43) is satisfied, or in the sense that both \mathfrak{A} and Ai_ϵ are close enough to $\delta(H_y(y) - h)$. Then, (5.60) becomes

$$\mu \approx \text{Ai}_\epsilon(H_y(y) - h) \delta(\tau(q) - \tau). \quad (5.62)$$

Because y and z are connected by a linear transformation, which preserves μ , one can also consider (5.62) as the symplectic measure for the whole transformation $r \leftrightarrow z$:

$$\mu = \text{Ai}_\epsilon(H_z(z) - h) \delta(\tau(z) - \tau). \quad (5.63)$$

At $\epsilon \rightarrow 0$, when the transformation $z \mapsto r$ is linear,³⁴ (5.63) yields $\mu(r, z) = \delta(H_z(z) - h) \delta(\tau(z) - \tau)$, in agreement with (3.57). Application of (5.63) in the

³⁴By (5.27), zero ϵ corresponds to inflection points on the dispersion curve, i.e. locally straight rays. Thus, the ray-aligned coordinates r are locally identical to y .

general case requires that one expresses ε as a function of \mathbf{z} (as opposed to q). Since our calculation is valid only for small q , multiple coordinate charts would have to be introduced and properly merged for this, which is inconvenient. However, as long as ε is smooth (or indistinguishable from zero), one can use (5.16) to extrapolate it to all \mathbf{z} via

$$\varepsilon = (\mathbf{J}\partial_{\mathbf{z}}H_{\mathbf{z}}) \cdot (\partial_{\mathbf{z}}^2 H_{\mathbf{z}})(\mathbf{J}\partial_{\mathbf{z}}H_{\mathbf{z}}). \quad (5.64)$$

This ensures that ε is close to the desired value $\varepsilon = -V^3 p''$ in the ray vicinity, while far from the ray, our result is not expected to be applicable in any case. Then, one can readily transform the symbols of operators with

$$O_{\mathbf{z}}(\mathbf{z}) \approx \int \text{Ai}_{\varepsilon}(H_{\mathbf{z}}(\mathbf{z}) - \tilde{h}) O_r(\tau(\mathbf{z}), \tilde{h}) d\tilde{h}, \quad (5.65)$$

i.e. $O_{\mathbf{z}}(\mathbf{z})$ is the Airy transform (Widder 1979) of O_r . Equation (5.65) is one of the key new results of this paper. As will be seen in the following, it is an important theorem that makes MGO a practical theory.

Assuming the symplectic scale of O_r is of order R , one can take this integral using (C2). This leads to $O_{\mathbf{z}}(\mathbf{z}) \approx O_r(\tau(\mathbf{z}), H_{\mathbf{z}}(\mathbf{z})) + \mathcal{O}(\varepsilon \partial_{\tilde{h}}^3 O_r)$. Based on (5.64), the second term on the right-hand side can be estimated as $\varepsilon \partial_{\tilde{h}}^3 O_r \sim O_r/R^4$. Thus,

$$O_{\mathbf{z}}(\mathbf{z}) = O_r(\mathbf{r}(\mathbf{z})) + \mathcal{O}(\varepsilon^2). \quad (5.66)$$

The importance of this result is in that the approximate equality $O_{\mathbf{z}}(\mathbf{z}) \approx O_r(\mathbf{r}(\mathbf{z}))$ is satisfied with accuracy not $\mathcal{O}(\varepsilon)$ but $\mathcal{O}(\varepsilon^2)$, which will be negligible for our purposes (§ 6). This is another new key result of this paper.³⁵ Keep in mind, though, that (5.66) is invalid for small-scale symbols, such as the delta-shaped symbols discussed in § 6.2. In that case, the more general (5.65) should be used instead.

6. MGO for scalar waves

6.1. Dynamics in the τ -space

Armed with (5.66), we are now ready to formulate the MGO envelope equation. Let us return to our original wave equation (4.16), which accounts for dissipation, and rewrite it in the τ -representation:

$$(\hat{H}_{\hat{\tau}} + i\hat{\Gamma}_{\hat{\tau}})\psi_{\hat{\tau}} = 0. \quad (6.1)$$

(As earlier, we assume that $\psi_{\hat{\tau}}$ is a scalar function, but see § 7.2 for a more general case.) Since $\hat{H}_{\hat{\tau}} = \hat{h}_{\hat{\tau}} = -i\partial_{\tau}$, this leads to

$$(-\partial_{\tau} + \hat{\Gamma}_{\hat{\tau}})\psi_{\hat{\tau}} = 0. \quad (6.2)$$

Assuming that $\Gamma_r = \mathcal{O}(\varepsilon)$, the operator $\hat{\Gamma}_{\hat{\tau}}$ can be approximated simply by its symbol on the reference ray, $\Gamma_r(\tau)$. Then, (6.2) becomes

$$\partial_{\tau}\psi_{\hat{\tau}} = \Gamma_r(\tau)\psi_{\hat{\tau}}. \quad (6.3)$$

³⁵A related, not physicist-friendly, treatise of this subject can be found in (Silva 2007), but it does not directly apply to the delta-shaped M -waves with which we deal here.

Assuming that Γ_r is sufficiently smooth, (5.66) predicts that $\Gamma_r(\mathbf{r}) \approx \Gamma_z(\mathbf{z}(\mathbf{r}))$, which is a known function when the dispersion operator is given. Then, (6.3) can be readily integrated,

$$\psi_{\hat{\tau}} = a \exp\left(\int d\tau \Gamma_r(\tau)\right), \quad a = \text{const.}, \quad (6.4)$$

and, unlike in the GO (4.12), this solution exhibits no singularities. Also notably, the ray equations in the \mathbf{r} -representation are particularly simple:

$$\dot{\hat{\tau}} = 1, \quad \dot{\hat{h}} = 0. \quad (6.5)$$

A generalisation to vector waves is discussed in § 7. One can also generalise the MGO envelope equation to quasioptical beams that experience transverse diffraction. This can be done similarly to how quasioptical equations were derived by Dodin *et al.* (2019) and Yanagihara, Dodin & Kubo (2019a,b) for conventional GO. However, presenting these calculations in detail would require introducing additional cumbersome machinery that is not specific to MGO, so we postpone discussing this subject until further publications.

6.2. Mapping solutions to the x -space

Mapping solutions to the x -space is generally the most challenging part in MGO applications (Donnelly *et al.* 2021; Højlund Marholt *et al.* 2024; Lopez, Højlund & Sensius 2024). However, knowing the field at all x is usually unnecessary. One might not even need to leave the τ -space except for initialising the field and also calculating the final state, and those mappings may be easy to do.³⁶ In other cases, one can proceed as follows.

6.2.1. General case

To calculate how the wave evolves in the original x -space, one needs to remap the τ -representation to the x -representation:

$$\psi_{\hat{x}}(x) = \int d\tau M_{\hat{x} \leftarrow \hat{\tau}}(x, \tau) \psi_{\hat{\tau}}(\tau). \quad (6.6)$$

Since $\hat{M}_{\hat{x} \leftarrow \hat{\tau}} = \hat{M}_{\hat{x} \leftarrow \hat{q}} \hat{M}_{\hat{q} \leftarrow \hat{\tau}}$, this leads to

$$\psi_{\hat{x}}(x) = \int dq \frac{e^{i\vartheta(q)}}{\sqrt{V(q)}} M_{\hat{x} \leftarrow \hat{q}}(x, q) \psi_{\hat{\tau}}(\tau(q)), \quad (6.7)$$

where $\vartheta(q) = \int_0^q p(\tilde{q}) d\tilde{q}$. Since $M_{\hat{x} \leftarrow \hat{q}}$ consists of a phase-space shift (§ 3.5.1) and an LST (§ 3.5.3), the integrand in (6.7) is easy to write explicitly in any given case.

How to approximate such an integral depends on the problem and can be done using various techniques, with accuracy requirements determined by specific applications. For example, stationary-phase or steepest-descent methods might suffice, or one might be able to reduce (6.7) to a known representation of a special function.

³⁶It may be worth drawing a parallel with quantum computing, where the cost of outputting the computation results can be a significant fraction of the total computation cost. Because of this, one usually aims to output ‘expectation values’ rather than the field profiles *per se* in quantum simulations, including simulations of linear waves (Dodin & Startsev 2021; Novikau, Dodin & Startsev 2022; Novikau, Startsev & Dodin 2023). We recommend adhering to this principle in MGO as well.

In particular, to the extent that (6.7) can be expressed through the Airy function (or, equivalently, Bessel functions of order 1/3), it is similar to previous results (Langer 1937; Chester, Friedman & Ursell 1957). Similar integrals have been studied also in the context of semiclassical calculations in atomic and molecular physics; for example, see Kay (1994b).

Remember also that the above-mentioned result is derived under the assumption that only small values of q contribute to the integral (6.7). This allows one to derive local approximations $\psi_{\hat{x}}$ using different definitions of (q, p) for different regions of interest. A global representation is not necessarily desirable for practical applications. When it is, though, it can be constructed in one of two ways.

One way to do this is to numerically find the type-1 generating function $\Theta(x, \tau)$ of the transformation $(x, k) \mapsto (\tau, h)$ using

$$k = \partial_x \Theta(x, \tau), \quad h = -\partial_\tau \Theta(x, \tau), \quad d\Theta = k dx - h d\tau, \quad (6.8)$$

and then invoke (6.6) in combination with the approximate formula (4.40). For the specific coordinates that we assume here, the latter takes the form

$$M_{\hat{x} \leftrightarrow \hat{\tau}}(x, \tau) = \sqrt{\frac{\Theta_{x\tau}}{2\pi}} e^{i\Theta}. \quad (6.9)$$

This formula is also useful for determining the quantisation condition for periodic orbits, as discussed in Appendix D.

Alternatively, one may notice that $\psi_{\hat{x}}$ *per se* is rarely needed in applications. Instead, of interest are bilinear functionals of $\psi_{\hat{x}}$ and $\psi_{\hat{x}}^*$, such as the energy density or the dissipated power. Any such functional can be expressed through the field's Wigner function.³⁷ By (2.43) and (6.4), the Wigner function of $\psi_{\hat{\tau}}$ in the r -representation is

$$W_r(\tau, h) \approx W_0(\tau)\delta(h), \quad W_0(\tau) = |a|^2 \exp\left(2 \int d\tau \Gamma_r(\tau)\right). \quad (6.10)$$

By (3.41), its z -representation is $W_z(z) = \int d\tau dh W_r(\tau, h) \mu(r, z)$. This leads to a remarkably concise expression for W_z :

$$W_z(z) = W_0(\tau(z)) \text{Ai}_\varepsilon(H_z(z)). \quad (6.11)$$

6.2.2. Eigenwaves

Let us now consider an important special case when dissipation is negligible, at least locally ($\hat{\Gamma} = 0$), and a wave satisfies $\hat{H}_{\hat{\tau}}\psi_{\hat{\tau}} \approx 0$. Such a wave can be understood as an eigenstate of \hat{H} corresponding to the zero eigenvalue, $h = 0$; i.e. $|\psi\rangle = a_h |\epsilon_{\hat{h}}(0)\rangle$, where $a_h \equiv \sqrt{2\pi}a$ is some constant amplitude. Then, using (2.17), one obtains

$$\psi_{\hat{\tau}}(\tau) = a_h \langle \epsilon_{\hat{\tau}}(\tau) | \epsilon_{\hat{h}}(0) \rangle = a = \text{const.}, \quad (6.12)$$

³⁷See Dodin (2022, Appendix B.1) for the general formula. In Appendix E of the present paper, we explicitly show how to derive this form for the dissipation power in particular. In fact, any reasonably general quasilinear calculations are expressed more naturally through Wigner functions than through fields *per se* (Dodin 2022, 2024). This is similar to the situation in quantum mechanics, where expectation values, as opposed to wavefunctions, are all that matter.

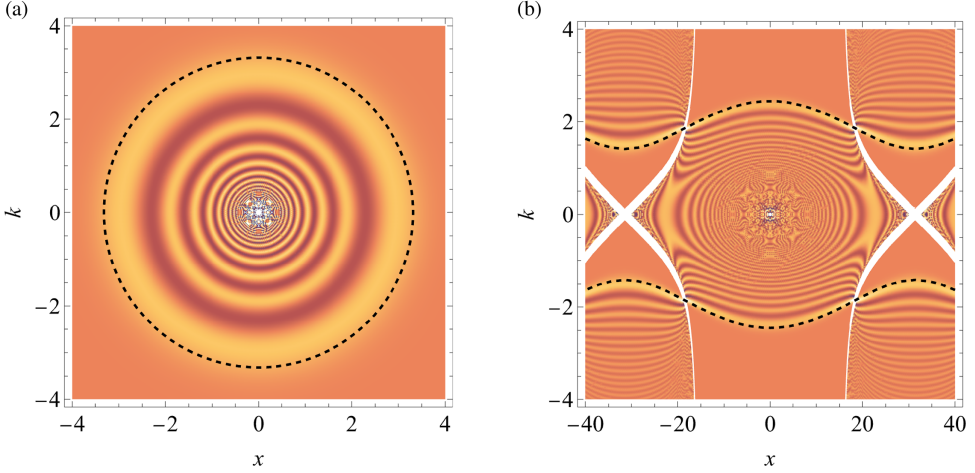


FIGURE 3. Approximation (6.16), with (5.64) for ε , for the Wigner functions W_z of fields satisfying (4.16): (a) $H_z(x, k) = (x^2 + k^2 - 11)/2$, which corresponds to a QHO with $n = 5$; (b) $H_z(x, k) = -2 - \cos(x/10) + k^2/2$. The intensity denotes the magnitude of W_z (arbitrary units). The white areas in panel (b) correspond to $\varepsilon \approx 0$. The approximation is accurate near the ray trajectories given by $H_z(x, k) = 0$ (dashed). In panel (b), there are two such trajectories, one for each sign of k . Strictly speaking, the Wigner function shown corresponds to a standing wave, to which both these branches contribute equally. Otherwise, although still applicable locally, (6.16) cannot be continued across the region where the contributions of the two disconnected branches are comparable (i.e. at $|k| \lesssim 1$ in this case).

which, of course, can also be obtained directly from (6.3). For $\psi_{\hat{x}}$, (6.7) readily yields

$$\psi_{\hat{x}} = \psi_{\hat{\tau}} \int dq \frac{e^{i\vartheta(q)}}{\sqrt{V(q)}} M_{\hat{x} \leftarrow \hat{q}}(x, q). \quad (6.13)$$

In the special case when $q \equiv x$, one has $M_{\hat{x} \leftarrow \hat{q}}(x, q) = \delta(x - q)$, so the standard WKB solution is obtained:

$$\psi_{\hat{x}}(x) = \psi_{\hat{\tau}} \frac{e^{i\vartheta(x)}}{\sqrt{V(x)}}. \quad (6.14)$$

More generally, (6.13) represents, expectedly, the same WKB solution in the tangent space, which is then mapped to the x -space using a linear MT. One can approach the integral (6.13) in the same way as discussed in § 6.2.1.

Alternatively, one can use (6.10) and (6.11), which, in this case, become

$$W_r(\tau, h) = W_0 \delta(h), \quad (6.15)$$

$$W_z(\mathbf{z}) \approx W_0 \text{Ai}_\varepsilon(H_z(\mathbf{z})), \quad (6.16)$$

with $W_0 = |a|^2 = \text{const}$. Examples of such $W_z(\mathbf{z})$ are presented in figure 3 for rays in a quadratic potential, which corresponds to a QHO, and a sinusoidal potential. Similar patterns are also seen in direct wave simulations; for example, see Zhu & Dodin (2021, figure 3a2) or Weinbub & Ferry (2018, figure 9). Note that these complicated patterns are merely a result of mapping the Wigner function to the original coordinates \mathbf{z} , as opposed to the natural coordinates \mathbf{r} , where the Wigner

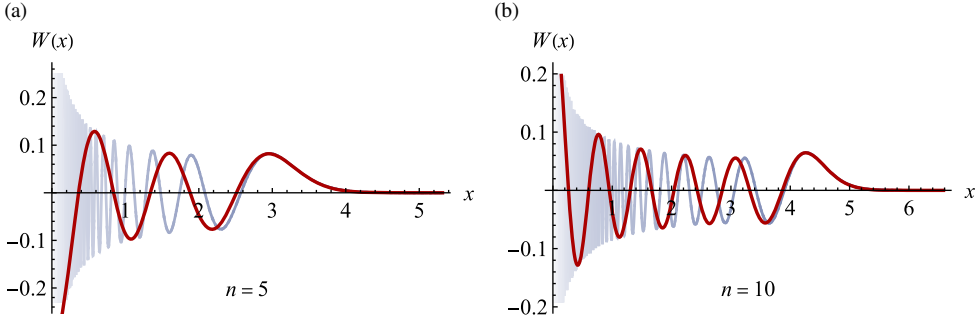


FIGURE 4. Wigner functions $W(x) \equiv W_z(x, k=0)$ of a QHO, $H_z = (x^2 + k^2)/2 - (n + 1/2)$: (a) $n = 5$; (b) $n = 10$. Red – exact analytical result (6.17); blue – approximation (6.16), with $W_0 = (2\pi)^{-1}$. The latter is valid only close to the ray (dashed curve in figure 3a), which, at $k = 0$ considered here, corresponds to the reflection points $x = \sqrt{2n + 1}$.

functions have a simple form (6.15). Also note that (6.15) is more precise than (6.16) in that it does not rely on approximations of the MT and is valid at all z . In contrast, (6.16) is accurate only in the vicinity of the ray trajectories, where H_z is small. This is illustrated in figure 4, which shows a comparison of (6.16) with the exact analytical result for a QHO, $H_z(x, k) = (x^2 + k^2)/2 - (n + 1/2)$ (Mostowski & Pietraszewicz 2021):

$$W_z(x, k) = \frac{(-1)^n}{\pi} e^{-\xi^2} L_n(2\xi), \quad \xi \doteq x^2 + k^2. \quad (6.17)$$

Here, n is a non-negative integer that corresponds to the state number, L_n is n th Laguerre polynomial, and the standard normalisation $a_h = 1$ is assumed, which corresponds to $W_0 = (2\pi)^{-1}$.

To the extent that Ai_ε can be replaced with the delta function, (6.16) yields

$$|\psi_{\hat{x}}(x)|^2 = \int dk W_z(x, k) \approx W_0 \int dk \delta(H_z(x, k)) = \frac{W_0}{|v_g|}, \quad (6.18)$$

where $v_g \doteq \partial_k H_z$ and the derivative is evaluated on the dispersion curve at a given x and the corresponding k on the dispersion curve. This coincides with the result that flows from the action conservation for stationary waves (see, e.g. Dodin 2022, §7.3.2),

$$\partial_x (v_g |\psi_{\hat{x}}|^2) = 0, \quad (6.19)$$

and $|\psi_{\hat{x}}|^2$ depends on x only due to inhomogeneity of v_g , i.e. slowly. If a wave experiences reflection, one might want to account for the multiple branches, i.e. two or more values of k at which the delta-function argument in (6.18) turns to zero at given x .³⁸ Then,

$$|\psi_{\hat{x}}(x)|^2 = \sum_i \frac{W_0}{|v_{g_i}|}, \quad (6.20)$$

³⁸Whether to account for one or multiple branches depends on the problem that one seeks to solve. For example, when resonant interactions with particles or other waves are considered, the resonance conditions are typically satisfied only for one branch. Then, the other branches are irrelevant and can be ignored.

where the summation is taken over said branches. Like in the previous case, the dependence on x (due to inhomogeneity of $|v_{\mathbf{g}}|i$) is slow in this case. This is because the assumed delta approximation for W_z ignores the interference between multiple branches; i.e. (6.20) can be understood as $|\psi_{\hat{x}}(x)|^2$ averaged over their relative phases.

One can also integrate the actual Wigner function (6.16) without resorting to the delta approximation, at least as long as the parabolic approximation (5.29) holds. For example, in the simplest case when $(q, p) \equiv (x, k)$, one readily obtains that

$$|\psi_{\hat{x}}(x)|^2 = W_0 \int dk \text{Ai}_{\varepsilon} \left(vk + \frac{vx^2}{2R} \right) = \frac{W_0}{v}, \quad (6.21)$$

which is in agreement with (6.18). More generally, when (q, p) are connected with (x, k) via (3.50) (we assume $z_0 = 0$ for brevity), one can write

$$H_z(z) = \underbrace{-\frac{v}{B} \left(x + \frac{RD^2}{2B} \right)}_{U(x)} + \frac{vB^2}{2R} \underbrace{\left(k + \frac{Ax}{B} + \frac{RD}{B^2} \right)^2}_K, \quad (6.22)$$

where we used (3.51). Then, using (Vallée *et al.* 1997)

$$\int_{-\infty}^{\infty} \text{Ai}(X + P^2) dP = 2^{2/3} \pi \text{Ai}^2(X/2^{2/3}) \quad (6.23)$$

and (5.24), one finds that

$$\begin{aligned} |\psi_{\hat{x}}(x)|^2 &= W_0 \int dK \text{Ai}_{\varepsilon} \left(U(x) + \frac{vB^2}{2R} K^2 \right) \\ &= W_0 \frac{\pi v}{2^{1/3}|B|} \text{Ai}_{\varepsilon}^2 \left(\frac{U(x)}{2^{2/3}} \right). \end{aligned} \quad (6.24)$$

In particular, the case when $B = 1$ and $D = 0$ corresponds to having a reflection point at $x = 0$, when the y -space is rotated by $\pi/2$ relative to the z -space. Then, one obtains

$$|\psi_{\hat{x}}(x)|^2 = 2\pi W_0 \frac{2^{2/3}}{|R|^{1/3}} \text{Ai}^2(-\sigma_{\varepsilon}|2R|^{1/3}x), \quad (6.25)$$

which corresponds to approximating (4.16) with the (appropriately rescaled) Airy equation near the reflection point. Although this is a commonly known result, a comparison of (6.25) with the exact solutions for a QHO is shown in figure 5 for completeness.

Let us also consider the opposite limit, that is, $B \rightarrow 0$ and $D \rightarrow 1$. One might expect to recover (6.21) in this case, yet (6.24) leads to a different result. Assuming the notation $\xi \doteq |R|^{2/3}/(2^{1/3}|B|) \gg 1$, it can be written as follows:

$$|\psi_{\hat{x}}(x)|^2 \rightarrow \frac{2W_0}{v} 2\pi\xi \text{Ai}^2(-\xi^2) \approx \frac{2W_0}{v} \left(1 + \sin \frac{4\xi^3}{3} \right). \quad (6.26)$$

This is understood as follows. Notice that now the problem is posed such that $|\psi_{\hat{x}}(x)|^2$ automatically includes the contributions from two branches, that is, the

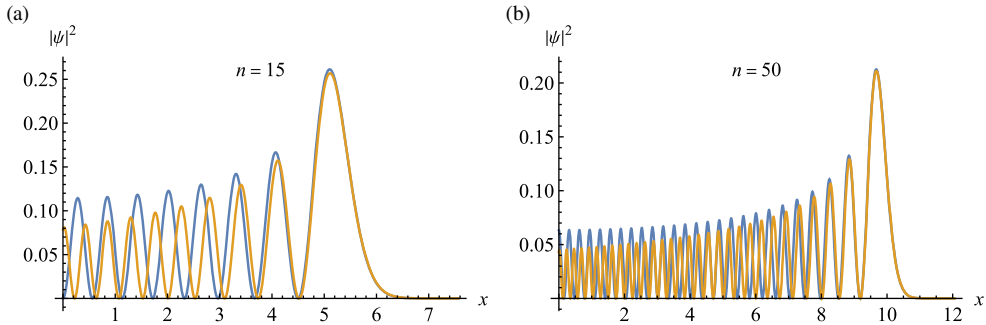


FIGURE 5. Normalised eigenstates $|\psi|^2 \equiv |\psi_{\hat{x}}(x)|^2$ of a QHO for two sample states: (a) $n = 15$ and (b) $n = 50$. Blue – exact analytical result; orange – approximation (6.25), with $W_0 = (2\pi)^{-1}$ and $R = \sqrt{2n+1}$. Unlike in (6.21)–(6.26), which assume $z_0 = 0$ for simplicity, the x -axis origin here is at the centre of the potential well, as usual; i.e. to reproduce these plots, one should replace x with $x - R$ in (6.25).

two roots $k(x)$ of the parabolic Hamiltonian (6.22). At $B \rightarrow 0$, $D \rightarrow 1$ and $x \rightarrow 0$, one of these roots approaches zero, while the other one tends to infinity, $k \approx -2R/B^2 \rightarrow -\infty$. To the extent that the interference with this second branch can be ignored, one can average out the second term in (6.26) to zero and obtain a result consistent with (6.20), $|\psi_{\hat{x}}(x)|^2 \rightarrow 2W_0/v$. Likewise, if the second branch is ignored entirely,³⁹ then

$$|\psi_{\hat{x}}(x)|^2 \rightarrow W_0/v, \quad (6.27)$$

as expected. This can equivalently be done by replacing the function Ai^2 in (6.26) with $|\text{Ai} + i\text{Gi}|^2/4$, where Gi is the Scorer function, to isolate the contribution from a single branch (Lopez 2024). The asymptotic form of $|\text{Ai} + i\text{Gi}|$ for large ξ then gives (6.27).

7. MGO for vector waves

7.1. Basic equations

Let us now consider the case when $\boldsymbol{\psi}_{\hat{\tau}}$ is an N -dimensional field,⁴⁰ and, accordingly, $\hat{\mathbf{H}}_{\hat{\tau}}$ and $\hat{\mathbf{\Gamma}}_{\hat{\tau}}$ are $N \times N$ matrices of operators (hence, the bold font),

$$(\hat{\mathbf{H}}_{\hat{\tau}} + i\hat{\mathbf{\Gamma}}_{\hat{\tau}}) |\boldsymbol{\psi}\rangle = 0. \quad (7.1)$$

In this case, one cannot adopt $\hat{h}_{\hat{\tau}}$ equal to $\hat{\mathbf{H}}_{\hat{\tau}}$, because $\hat{h}_{\hat{\tau}} = -i\partial_{\tau}$ must be a one-dimensional operator. (One can convert any Hermitian combination of the elements of $\hat{\mathbf{H}}_{\hat{\tau}}$ into $-i\partial_{\tau}$, but not all of them at the same time.) Instead, we proceed as follows. Let us Weyl-expand (7.1) by analogy with (4.8), using the smallness of h near the reference ray:

$$(\mathbf{H}_r(\tau) + i\mathbf{\Gamma}_r(\tau))\boldsymbol{\psi}_{\hat{\tau}} - i\mathbf{V}(\tau)\partial_{\tau}\boldsymbol{\psi}_{\hat{\tau}} - \frac{i}{2}\mathbf{V}'(\tau)\boldsymbol{\psi}_{\hat{\tau}} \approx 0. \quad (7.2)$$

³⁹See the previous footnote.

⁴⁰The dimension of $\boldsymbol{\psi}_{\hat{\tau}}$ is generally unrelated to the dimension of the coordinate space.

Here, the reference ray is yet to be specified and so is \hat{h} . Also, the prime denotes a derivative with respect to the argument, $\mathbf{H}_r(\tau)$ is the symbol \mathbf{H}_r on the reference ray,⁴¹ $\mathbf{\Gamma}_r(\tau)$ is the symbol $\mathbf{\Gamma}_r$ on the reference ray and

$$\mathbf{V}(\tau) \doteq (\partial_h \mathbf{H}_r(\tau, h))_{h=0}. \quad (7.3)$$

As we have already mentioned in §6.1, this approach also allows including transverse diffraction, following the same procedure as used by Dodin *et al.* (2019) and Yanagihara *et al.* (2019a,b) for conventional GO, but we leave this generalisation to future work.

7.2. Single mode

Since $\mathbf{\Gamma}_r$ and ∂_τ are order- ϵ in (7.2), the term $\mathbf{H}_r \boldsymbol{\psi}_{\hat{z}}$ must be order- ϵ as well. This means that $\boldsymbol{\psi}_{\hat{z}}$ is close to a local normalised eigenvector $\boldsymbol{\eta}$ of \mathbf{H}_r that corresponds to some small eigenvalue Λ (which is real, because \mathbf{H}_r is Hermitian); i.e.

$$\mathbf{H}_r \boldsymbol{\eta} = \Lambda \boldsymbol{\eta}, \quad \boldsymbol{\eta}^\dagger \boldsymbol{\eta} = 1, \quad \Lambda = \mathcal{O}(\epsilon). \quad (7.4)$$

Let us assume, until §7.3, that \mathbf{H}_z is non-degenerate, so that Λ is its only small eigenvalue. Then, such $\boldsymbol{\psi}_{\hat{z}}$ can be represented as

$$\boldsymbol{\psi}_{\hat{z}} = \eta a + \bar{\boldsymbol{\Psi}}, \quad \bar{\boldsymbol{\Psi}} = \mathcal{O}(\epsilon), \quad \boldsymbol{\eta}^\dagger \bar{\boldsymbol{\Psi}} = 0, \quad (7.5)$$

where $a = a(\tau)$ is a scalar amplitude and $\bar{\boldsymbol{\Psi}}$ is the (small) component of $\boldsymbol{\psi}_{\hat{z}}$ transverse to $\boldsymbol{\eta} = \boldsymbol{\eta}(\tau)$. This means that the wave consists mostly of a single mode⁴² with polarisation $\boldsymbol{\eta}$, and the non-zero contribution $\bar{\boldsymbol{\Psi}}$ from the other modes is only due to the fact that the wave evolves along the ray and therefore is not entirely monochromatic in τ .

Let us multiply (7.2) by $\boldsymbol{\eta}^\dagger$ from the left. After omitting $\mathcal{O}(\epsilon^2)$ terms, this gives⁴³

$$(\Lambda + i \underbrace{\boldsymbol{\eta}^\dagger \mathbf{\Gamma}_r \boldsymbol{\eta}}_r) a - i \boldsymbol{\eta}^\dagger \mathbf{V} \boldsymbol{\eta}' a - \frac{i}{2} \boldsymbol{\eta}^\dagger \mathbf{V}' \boldsymbol{\eta} a - i \underbrace{\boldsymbol{\eta}^\dagger \mathbf{V} \boldsymbol{\eta}}_v \partial_\tau a = 0, \quad (7.6)$$

or, equivalently,

$$(\Lambda - U + i\Gamma) a - i\mathcal{V} \partial_\tau a - \frac{i}{2} \mathcal{V}' a = 0, \quad U \doteq \text{im}((\boldsymbol{\eta}^\dagger)' \mathbf{V} \boldsymbol{\eta}), \quad (7.7)$$

where we used

$$\begin{aligned} -i \boldsymbol{\eta}^\dagger \mathbf{V} \boldsymbol{\eta}' - \frac{i}{2} \boldsymbol{\eta}^\dagger \mathbf{V}' \boldsymbol{\eta} &= -i \boldsymbol{\eta}^\dagger \mathbf{V} \boldsymbol{\eta}' - \frac{i}{2} \partial_\tau (\boldsymbol{\eta}^\dagger \mathbf{V} \boldsymbol{\eta}) + \frac{i}{2} (\boldsymbol{\eta}^\dagger)' \mathbf{V} \boldsymbol{\eta} + \frac{i}{2} \boldsymbol{\eta}^\dagger \mathbf{V} \boldsymbol{\eta}' \\ &= -\frac{i}{2} \mathcal{V}' - \frac{i}{2} (\boldsymbol{\eta}^\dagger \mathbf{V} \boldsymbol{\eta}' - (\boldsymbol{\eta}^\dagger)' \mathbf{V} \boldsymbol{\eta}) \\ &= -\frac{i}{2} \mathcal{V}' - U. \end{aligned} \quad (7.8)$$

⁴¹As a reminder, the symbol $\mathbf{A} \doteq \text{ymb} \hat{\mathbf{A}}$ of a given matrix operator $\hat{\mathbf{A}}$ is simply the matrix composed of the symbols of the original matrix elements, $\text{ymb} \hat{\mathbf{A}}_{ab}$.

⁴²By modes, we mean monochromatic waves that would have been exact solutions at $\partial_\tau = 0$.

⁴³As usual, the prime denotes a derivative with respect to the argument, i.e. τ . Also, $\partial_\tau \equiv a'$ here, but would be a partial derivative for quasioptical beams in a multi-dimensional space.

Also note that $\mathcal{V} \equiv \boldsymbol{\eta}^\dagger (\partial_h \mathbf{H}_r) \boldsymbol{\eta}$ can be written as

$$\begin{aligned}
 \mathcal{V} &= \frac{\partial}{\partial h} (\boldsymbol{\eta}^\dagger \mathbf{H}_r \boldsymbol{\eta}) - \frac{\partial \boldsymbol{\eta}^\dagger}{\partial h} \mathbf{H}_r \boldsymbol{\eta} - \boldsymbol{\eta}^\dagger \mathbf{H}_r \frac{\partial \boldsymbol{\eta}}{\partial h} \\
 &= \frac{\partial \Lambda}{\partial h} - \frac{\partial \boldsymbol{\eta}^\dagger}{\partial h} \Lambda \boldsymbol{\eta} - \boldsymbol{\eta}^\dagger \Lambda \frac{\partial \boldsymbol{\eta}}{\partial h} \\
 &= \frac{\partial \Lambda}{\partial h} - \frac{\partial (\boldsymbol{\eta}^\dagger \boldsymbol{\eta})}{\partial h} \Lambda \\
 &= \frac{\partial \Lambda}{\partial h}, \tag{7.9}
 \end{aligned}$$

because $\boldsymbol{\eta}^\dagger \boldsymbol{\eta} \equiv 1$. Then, it is convenient to choose \hat{h} such that $h = \Lambda$, so (7.9) yields $\mathcal{V} = 1$.

Now, let us finally specify the reference ray. One option is to choose it such that $\Lambda = 0$ on this ray, as most commonly done in GO (Stix 1992). Then, (7.7) becomes

$$\partial_\tau a = (\Gamma + iU)a, \quad \partial_\tau |a|^2 = 2\Gamma |a|^2, \tag{7.10}$$

which is the MGO counterpart of the GO (4.12). An advantage of this approach is the simplicity of the ray Hamiltonian, Λ . A disadvantage is that the term U causes the phase of a to evolve. Over large time, this can make a a rapidly oscillating function (across the rays in multi-dimensional problems), potentially undermining the assumed ordering and thus rendering (7.10) inaccurate.

Alternatively, one can choose the reference ray such that $\Lambda - U = 0$, so (7.7) becomes

$$\partial_\tau a = \Gamma a, \quad \partial_\tau |a|^2 = 2\Gamma |a|^2. \tag{7.11}$$

The advantage of this approach is that the phase of a remains fixed, so a is the true envelope and is more likely to remain smooth. However, since the ray Hamiltonian is now $\Lambda - U$, the ray equations become more complicated. The resulting reference ray deviates from the conventional GO trajectory governed by the Hamiltonian Λ . In GO, this deviation is called the spin-Hall effect. For further discussion, see Littlejohn & Flynn (1991), Bliokh *et al.* (2015), Ruiz & Dodin (2015), Ruiz & Dodin (2017a), Dodin *et al.* (2019) and Fu, Dodin & Qin (2023), and references therein.

Solutions of the vector (7.10) and (7.11) can be mapped to the x -representation just like described in § 6.2. Importantly, the fact that $\boldsymbol{\psi}_{\hat{\tau}}$ is a vector has nothing to do with M -waves, which are constructed based just on the scalar ray Hamiltonian Λ , and therefore remain scalar.

7.3. Mode conversion

Suppose now that \mathbf{H}_r is degenerate, so that more than one of its eigenvalues are $\mathcal{O}(\epsilon)$. This means that the wave can consist of multiple modes that can resonantly interact with each other, i.e. experience linear mode conversion. In this case, one can represent $\boldsymbol{\psi}_{\hat{\tau}}$ as

$$\boldsymbol{\psi}_{\hat{\tau}} = \boldsymbol{\Xi} \mathbf{a} + \bar{\boldsymbol{\Psi}}, \quad \bar{\boldsymbol{\Psi}} = \mathcal{O}(\epsilon), \tag{7.12}$$

where the (generally non-square) matrix $\boldsymbol{\Xi}$ has the orthonormal polarisation vectors $\boldsymbol{\eta}_i$ of the resonant modes (eigenvectors of \mathbf{H}_r) as its columns, \mathbf{a} is the vector that

characterises the amplitudes of these modes and $\bar{\Psi}$ is the nonresonant part of $\psi_{\hat{r}}$. For details, see Dodin *et al.* (2019), which discusses a similar parametrisation within GO.

The derivation of the envelope equation in this case is identical to that in § 7.2 up to replacing η with \mathcal{E} (and a with \mathbf{a}), and one obtains

$$(\mathbf{A} - \mathbf{U} + i\mathbf{\Gamma})\mathbf{a} - i\mathcal{V}\partial_{\tau}\mathbf{a} - \frac{i}{2}\mathcal{V}'\mathbf{a} = 0, \quad \mathbf{U} \doteq ((\mathcal{E}^{\dagger})'\mathcal{V}\mathcal{E})_{\Lambda}. \quad (7.13)$$

Here, the index Λ denotes the anti-Hermitian part and $\mathbf{A} = \mathcal{E}^{\dagger}\mathbf{H}_r\mathcal{E}$ is a diagonal matrix, $\mathbf{A} = \text{diag}\{\Lambda_1, \Lambda_2, \dots\}$, where Λ_i are the small eigenvalues of \mathbf{H}_r . Also, $\mathbf{\Gamma} = \mathcal{E}^{\dagger}\mathbf{\Gamma}_r\mathcal{E}$ and $\mathcal{V} = \mathcal{E}^{\dagger}\mathcal{V}\mathcal{E}$, so the elements of the latter, $\mathcal{V}_{ij} = \eta_i^{\dagger}(\partial_h\mathbf{H}_r)\eta_j$, can be expressed as

$$\begin{aligned} \mathcal{V}_{ij} &= \frac{\partial}{\partial h} (\eta_i^{\dagger}\mathbf{H}_r\eta_j) - \frac{\partial\eta_i^{\dagger}}{\partial h}\mathbf{H}_r\eta_j - \eta_i^{\dagger}\mathbf{H}_r\frac{\partial\eta_j}{\partial h} \\ &= \frac{\partial\Lambda_{ij}}{\partial h} - \frac{\partial\eta_i^{\dagger}}{\partial h}\Lambda_j\eta_j - \eta_i^{\dagger}\Lambda_i\frac{\partial\eta_j}{\partial h}. \end{aligned} \quad (7.14)$$

In the envelope equation, the coefficients \mathcal{V}_{ij} appear only in combination with $\partial_{\tau} = \mathcal{O}(\epsilon)$, so they are of interest only to the zeroth order in ϵ . Then, the last two terms in (7.14), which are proportional to small eigenvalues of \mathbf{H}_r , are negligible; i.e. $\mathcal{V} \approx \partial_h\mathbf{A}$. In other words, \mathcal{V} is approximately diagonal, so mode coupling is determined entirely by \mathbf{U} and $\mathbf{\Gamma}$. Also, if $\det\mathcal{V} \approx \prod_i\partial_h\Lambda_i$ is non-zero, i.e. all $\partial_h\Lambda_i$ are non-zero, then \mathcal{V} is invertible and one can rewrite (7.13) as

$$i\partial_{\tau}\mathbf{b} = (\mathcal{N}^{-1}(\mathbf{A} - \mathbf{U} + i\mathbf{\Gamma})\mathcal{N}^{-1} + \mathfrak{H})\mathbf{b}, \quad (7.15)$$

where $\mathbf{b} \doteq \mathcal{N}\mathbf{a}$, $\mathcal{N} = \mathcal{V}^{1/2}$ and⁴⁴

$$\begin{aligned} \mathfrak{H} &= -i\mathcal{N}(\mathcal{N}^{-1})' - \frac{i}{2}\mathcal{N}^{-1}(\mathcal{N}\mathcal{N}')\mathcal{N}^{-1} \\ &= i\mathcal{N}'\mathcal{N}^{-1} - \frac{i}{2}\mathcal{N}^{-1}\mathcal{N}'\mathcal{N}\mathcal{N}^{-1} - \frac{i}{2}\mathcal{N}^{-1}\mathcal{N}\mathcal{N}'\mathcal{N}^{-1} \\ &= i\mathcal{N}'\mathcal{N}^{-1} - \frac{i}{2}\mathcal{N}^{-1}\mathcal{N}' - \frac{i}{2}\mathcal{N}'\mathcal{N}^{-1} \\ &= \frac{i}{2}(\mathcal{N}'\mathcal{N}^{-1} - \mathcal{N}^{-1}\mathcal{N}'). \end{aligned} \quad (7.16)$$

In particular, at zero $\mathbf{\Gamma}$, the operator on the right-hand side of (7.15) is Hermitian, so $|\mathbf{b}|^2$ is conserved. This reflects conservation of the total action of resonant waves.

In a typical scenario involving two resonant modes (Tracy, Kaufman & Brizard 2003, 2014), the corresponding dispersion curves are roughly parallel to each other in the mode-conversion region due to the ‘level repulsion’. Using a proper variable transformation, the mode-conversion region usually can be put in the symmetric

⁴⁴The fact that \mathcal{V} is approximately diagonal is enough to prove that it is invertible, i.e. that \mathcal{N}^{-1} exists, but not necessarily enough for \mathfrak{H} to be negligible.

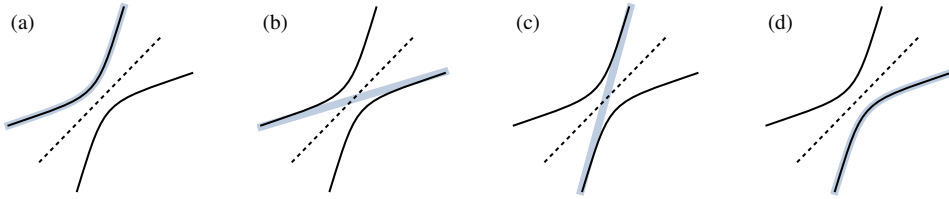


FIGURE 6. A typical symmetrised structure of the dispersion curves (black) of two resonant waves in the mode-conversion region in phase space \mathbf{z} . The blue curves indicate possible choices of the reference ray, depending on a problem that one needs to solve. The dashed line shows one of the symmetry axes.

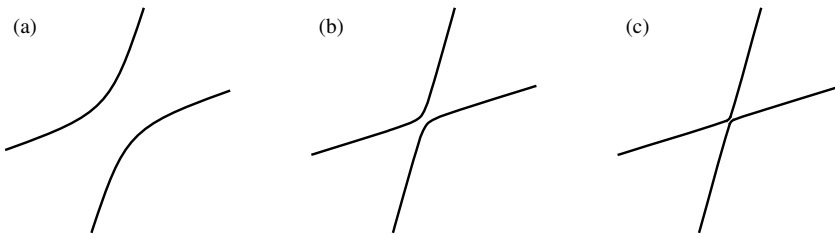


FIGURE 7. Schematic of how the dispersion curves of two resonance waves in the mode-conversion region in the \mathbf{z} -space transition from (a) smooth to (b) non-smooth as the exact resonance is approached.

form shown in [figure 6](#). If the reference ray is chosen such that, in the mode-conversion region, it propagates somewhat parallel to the axis of symmetry shown with the dashed line. In this case, both $\partial_h \Lambda_1$ and $\partial_h \Lambda_2$ are non-zero, so \mathcal{V} is invertible. Then, both (7.13) and (7.15) predict non-singular $\psi_{\hat{z}}$, regardless of the orientation of the dispersion curves in the original \mathbf{z} -space (i.e. MGO can handle mode conversion near cutoffs). The optimal choice of the Hamiltonian for the reference ray in this case depends on a problem that one needs to solve ([figure 6](#)). For example, if one is interested in how much energy is retained on the initial branch after the mode conversion, one might want to choose $h = \Lambda_1$ or $h = \Lambda_2$, depending on the initial conditions, so the reference ray simply follows one of the branches. Alternatively, if one is interested in how much energy is transmitted through the gap, one can choose $h = c_1 \Lambda_1 + c_2 \Lambda_2$, where the coefficients c_1 and c_2 are such that one of them vanishes at $\tau \rightarrow -\infty$ and the other one vanishes at $\tau \rightarrow +\infty$.

Keep in mind, however, that describing mode conversion using (7.13) and (7.15) may not be optimal when the gap between the dispersion curves is particularly small, resulting in large Λ' and X' ([figure 7](#)). In this case, representing $\psi_{\hat{z}}$ in the eigenvector basis, as in (7.12), may be inconvenient. Instead, one might want to solve the original equation (7.2), where the coefficients are smooth, unlike in (7.13) and (7.15). We will elaborate on this subject in a separate publication.

Solutions of (7.15) (or, if needed, (7.2)) can be mapped to the x -representation mostly like described in § 6.2. The only difference is that, in this case, the Wigner function of the wavefield becomes a Wigner matrix, which is the symbol of $(2\pi)^{-1} |\psi\rangle \langle \psi|$.

7.4. Example

As an illustration, let us consider a standard example (Tracy *et al.* 2014, Chapter 6) of mode conversion governed by the following differential equation for a two-dimensional vector field $\psi_{\hat{x}}$:

$$\begin{pmatrix} \frac{1}{2}(-i\partial_x - x) - i\gamma & \frac{\alpha}{2} \\ \frac{\alpha^*}{2} & \frac{1}{2}(-i\partial_x + x) - i\gamma \end{pmatrix} \psi_{\hat{x}} = 0. \quad (7.17)$$

Here, α is a complex constant and γ is a positive constant that introduces (positive) dissipation. This corresponds to

$$\hat{\mathbf{H}} = \begin{pmatrix} \frac{\hat{k} - \hat{x}}{2} & \frac{\alpha}{2} \\ \frac{\alpha^*}{2} & \frac{\hat{k} + \hat{x}}{2} \end{pmatrix} \quad (7.18)$$

and $\hat{\Gamma} = -\hat{\mathbf{1}}\gamma$. Then, the dispersion relation (without dissipation) has the form $k^2 - x^2 = |\alpha|^2$. Its solution has two branches, $k = \pm\sqrt{x^2 + |\alpha|^2}$, separated by a gap of size $2|\alpha|$ (red curves in figure 8), and the action flow on them is in the same direction, as seen from the fact that the corresponding $\partial_k \mathbf{H}(0, \mathbf{k})$ is a scalar matrix. The utility of this example is that it is analytically solvable and illustrates the typical features of mode conversion. (However, see Dodin, Ruiz & Kubo (2017) for situations that cannot be described within this model.) Also note that this problem is similar to the one with

$$\hat{\mathbf{H}} = \begin{pmatrix} -\frac{\hat{k} - \hat{x}}{2} & \frac{\alpha}{2} \\ \frac{\alpha^*}{2} & \frac{\hat{k} + \hat{x}}{2} \end{pmatrix}, \quad (7.19)$$

which is different only in that that the corresponding phase plot is rotated by 90° . In this case, the apex points become reflection points and the problem becomes akin to the one that governs the O–X conversion mentioned in § 1.1, but we postpone the discussion about this subject until a follow-up paper.

7.4.1. Approach 1

Suppose that the wave energy is on the lower branch at $x \rightarrow -\infty$ and we are interested in how much of it is transferred to the upper branch at $x \rightarrow +\infty$. Then, one can choose the reference ray as a line $k = x$ that asymptotically approaches the lower branch at $x \rightarrow -\infty$ and the upper branch at $x \rightarrow +\infty$ (blue line in figure 8a). Let us adopt

$$\hat{h} = \hat{k} - \hat{x}, \quad \hat{\tau} = (\hat{k} + \hat{x})/2, \quad (7.20)$$

which satisfy $[\hat{\tau}, \hat{h}] = i$, as needed. This leads to

$$\mathbf{H}_r = \begin{pmatrix} h/2 & \alpha/2 \\ \alpha^*/2 & \tau \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} 1/2 & 0 \\ 0 & 0 \end{pmatrix}. \quad (7.21)$$

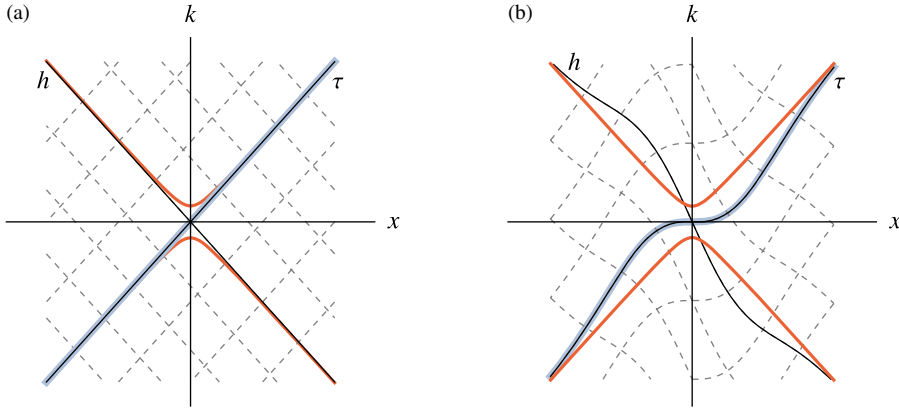


FIGURE 8. Dispersion curves (red) in the mode-conversion region with \hat{H} given by (7.18). The blue curves are the reference rays. The black lines are the axes τ and h , i.e. the curves corresponding to $h = 0$ and $\tau = 0$, respectively. The dashed lines show the corresponding coordinate grids at non-zero τ and h , as in: (a) (7.20); (b) (7.27).

Note that V is not invertible in this case and (7.2) yields

$$i\partial_\tau \psi_1 = -2i\gamma \psi_1 + \alpha \psi_2, \quad 0 = (\alpha^*/2)\psi_1 + (\tau - i\gamma)\psi_2. \quad (7.22)$$

Here, ψ_1 and ψ_2 are the two components of $\psi_{\hat{t}}$ and we used the fact that $h = 0$ on the reference ray. Eliminating ψ_2 using the second equation, one obtains a self-contained equation for ψ_1 :

$$\partial_\tau \psi_1 = \left(-2\gamma + \frac{i|\alpha|^2}{2(\tau - i\gamma)} \right) \psi_1 \rightarrow \frac{i|\alpha|^2}{2(\tau - i0)} \psi_1, \quad (7.23)$$

where we took the limit $\gamma \rightarrow 0$ for simplicity and $i0$ denotes an infinitesimal imaginary number with a positive imaginary part. This can be integrated as follows:

$$\ln T \equiv \ln \frac{\psi_1(+\infty)}{\psi_1(-\infty)} = \frac{i|\alpha|^2}{2} \int_{-\infty}^{\infty} \frac{1}{\tau - i0} d\tau = -\frac{\pi|\alpha|^2}{2}. \quad (7.24)$$

Then, the transmission coefficient for the amplitude is $T = e^{-\pi|\alpha|^2/2}$ and the transmission coefficient for the action flux $j_1 \doteq V_{11}|\psi_1|^2$ is, correspondingly,

$$|T|^2 = \frac{j_1(+\infty)}{j_1(-\infty)} = e^{-\pi|\alpha|^2}. \quad (7.25)$$

Expectedly, the larger the gap between the two branches of the dispersion relation, the less amount of wave action (energy, etc.) tunnels from one branch to the other.

7.4.2. Approach 2

In the simple example considered in Approach 1, the linear transformation (7.20) is sufficient globally, so calculation of ψ_1 does not *per se* require the upgraded version of MGO that we present in this paper. That said, notice that the solution is problematic at $\gamma \rightarrow 0$. According to (7.23), $\tau = 0$ is a singular point for ψ_1 in this

case, which is caused by non-invertibility of \mathbf{V} . Although not an issue for analytical calculations, this creates an inconvenience for numerical simulations. One can avoid this by bending the reference ray such that \mathbf{V} becomes invertible and then the new MGO machinery comes in handy.

For example, consider a reference ray in the form $k = x - f(x)$, where $f(x \rightarrow \pm\infty) = 0$. We assume $f(x) = xe^{-\beta x^2}$ with $\beta = 0.01$ (the specific f is not particularly important), then the ray looks like the blue curve in [figure 8\(b\)](#). One can adopt

$$\hat{h} = \hat{k} - \hat{x} + f(\hat{x}), \quad \hat{\tau} = (\hat{k} + \hat{x} + f(\hat{x}))/2, \quad (7.26)$$

which are easily seen to satisfy $[\hat{\tau}, \hat{h}] = i$. The inverse transformation is given by

$$\hat{x} = \hat{\tau} - \hat{h}/2, \quad \hat{k} = \hat{\tau} + \hat{h}/2 - f(\hat{\tau} - \hat{h}/2). \quad (7.27)$$

Then,

$$\mathbf{H}_r(\tau, h) \approx \begin{pmatrix} \frac{1}{2}(h - f(\tau - h/2)) & \frac{\alpha}{2} \\ \frac{\alpha^*}{2} & \tau - \frac{1}{2}f(\tau - h/2) \end{pmatrix}, \quad (7.28)$$

so, on the ray, where $h = 0$, one has

$$\mathbf{H}_r \approx \begin{pmatrix} -\frac{1}{2}f(\tau) & \frac{\alpha}{2} \\ \frac{\alpha^*}{2} & \tau - \frac{1}{2}f(\tau) \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} \frac{1}{2} + \frac{1}{4}f'(\tau) & 0 \\ 0 & \frac{1}{4}f'(\tau) \end{pmatrix} \quad (7.29)$$

and $\mathbf{V}' = \mathbf{1}f''(\tau)/4$. (The matrix \mathbf{V} still becomes non-invertible at $|\tau| \rightarrow \infty$, but this does not matter, because this model is needed only in the mode-conversion region, i.e. at relatively small τ .) Substituting these into [\(7.2\)](#), one obtains

$$\left(-\frac{f}{2} - i\gamma\right)\psi_1 + \frac{\alpha}{2}\psi_2 - \frac{i}{2}\left(1 + \frac{f'}{2}\right)\psi'_1 - \frac{i}{8}f''\psi_1 = 0, \quad (7.30a)$$

$$\frac{\alpha^*}{2}\psi_1 + \left(\tau - \frac{f}{2} - i\gamma\right)\psi_2 - \frac{i}{4}f'\psi'_2 - \frac{i}{8}f''\psi_2 = 0, \quad (7.30b)$$

where we omitted the argument τ for brevity. A numerical solution of [\(7.30\)](#) for $\gamma = 0$ is shown in [figure 9\(b\)](#). (For small non-zero γ , the solution looks similar up to dissipation, which gradually reduces j_1 much like in [figure 9a](#).) Note that the singularity is gone and $j_1(\tau)/j_1(-\infty)$ again asymptotes to the value given by [\(7.25\)](#), as it should. (Since the coordinates [\(7.26\)](#) coincide with [\(7.20\)](#) at large τ , the action fluxes should coincide too.) This illustrates the fact that MGO equations in curved phase-space coordinates are in agreement with the earlier theory and that MGO can be useful even in toy problems where global linear coordinate transformations are, formally, sufficient.

8. Wave-particle interactions in MGO

8.1. Generalised Cherenkov resonance

The MT formalism also suggests a way to rethink wave-particle interactions. Consider the Cherenkov condition of the wave-particle resonance:

$$\omega = kv_p, \quad (8.1)$$

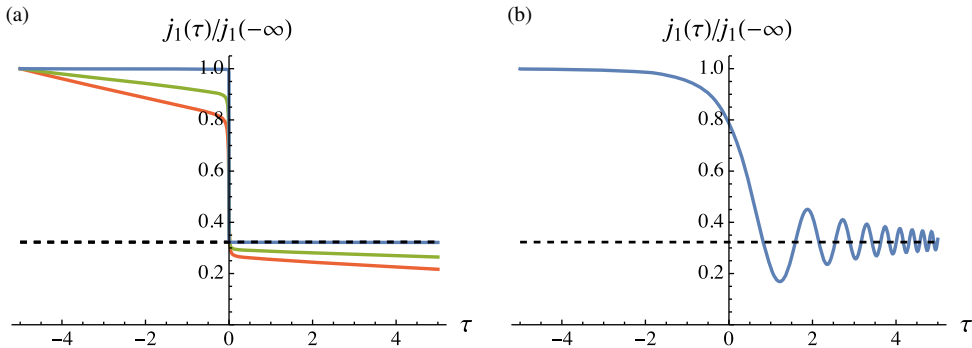


FIGURE 9. (a) Action flux $j_1 \doteq V_{11}|\psi_1|^2$ normalised to its initial value $j_1(-\infty)$, from the numerical solution of (7.23) for $\alpha = 0.6$ and $\gamma = 0.01$ (red), 0.005 (green), and 0.0001 (blue). Dashed is the analytical prediction for the limit at $\tau \rightarrow +\infty$ given by (7.25). (b) Same for (7.30) for $\gamma = 0$. In panel (a), $V_{11} \equiv 1/2$, so the plot is identical to that of $|\psi_1(\tau)/\psi_1(-\infty)|^2$. In panel (b), $V_{11} \equiv 1/2 + f'(\tau)/4$, which approaches $1/2$ only at $|\tau| \gtrsim 30$, so $j_1(\tau)/j_1(-\infty)$ and $|\psi_1(\tau)/\psi_1(-\infty)|^2$ asymptote to the same value, but the former saturates faster.

where ω is the wave frequency, k is the local wavenumber and v_p is the particle velocity. This condition is not symplectically invariant, but can be generalised within MGO, if one considers particles semiclassically.⁴⁵ Like any other waves, a semiclassical particle can be assigned phase-space variables (τ, h) . Suppose such a particle interacts with another (e.g. classical electromagnetic) wave that is eikonal in these variables too, even though this wave might not propagate along the curve $h = 0$. We assume this wave in the form

$$\psi_{\hat{\tau}}(t, \tau) = \Psi(t, \tau)e^{i\theta(t, \tau)}. \quad (8.2)$$

Here, Ψ is the envelope that is slow compared with the (real) phase θ and the time t can be considered as a parameter within our formalism. A particle can interact resonantly with such a wave if θ changes slowly, or not at all, in the particle frame:

$$0 \approx \dot{\theta} = \partial_t \theta + \dot{\tau} \partial_\tau \theta, \quad (8.3)$$

where the dot stands for d/dt .⁴⁶ Provided that the particle Hamiltonian and the wave Hamiltonian are normalised consistently, one has $\dot{\tau} = 1$ (cf. (6.5)). Then, assuming the standard definition $\omega \doteq -\partial_t \theta$, (8.3) leads to a metaplectic resonance condition:

$$\omega = \partial_\tau \theta. \quad (8.4)$$

(An alternative derivation and interpretation of this resonance condition is given in § 8.2.) Since $\dot{\tau} = 1$, all quantities here are entirely classical, i.e. \hbar is not involved. Also notice that the τ -axis can be aligned arbitrarily relative to the physical-space axis x . For example, a wave can resonantly interact with a particle even near a cutoff in

⁴⁵See Ruiz & Dodin (2017b) and Dodin & Fisch (2014) for how the Cherenkov condition emerges in generic wave-wave interactions and semiclassical interactions in particular.

⁴⁶Equation (8.3) subsumes (8.1) as a special case. Indeed, if the τ -axis is parallel to the x -axis, one has $\dot{\tau} \partial_\tau \theta = \dot{x} \partial_x \theta = kv_p$, where we used $k \doteq \partial_x \theta$ and $v_p \doteq \dot{x}$.

the x -space. This unifies Cherenkov acceleration with Fermi acceleration, which we define here broadly as particle acceleration caused by interaction with moving walls.

To better understand the condition (8.4), recall that $d\tau = d\phi/\Omega$. As a reminder, ϕ and Ω are the canonical phase and frequency of particle's OHO (§ 5.2),

$$\Omega = |\det(\partial_z^2 H_z)|^{1/2} \operatorname{sgn} \varepsilon, \quad (8.5)$$

where H_z is the particle Hamiltonian, z is the particle phase-space variables and $\varepsilon \doteq \ddot{z} \wedge \dot{z}$.⁴⁷ Then, (8.4) can be written as

$$\frac{\partial \theta}{\partial \phi} = \frac{\omega}{\Omega}. \quad (8.6)$$

For illustration, let us consider a special case when ω is a constant and the particle oscillates (non-relativistically) in a parabolic potential, so Ω is a constant too. Then, assuming (8.6) is satisfied globally, integrating this equation over ϕ from 0 to 2π gives

$$\frac{\Delta \theta_{\odot}}{2\pi} = \frac{\omega}{\Omega}, \quad (8.7)$$

where $\Delta \theta_{\odot}$ is the total change of θ on said interval. Since $m \doteq \Delta \theta_{\odot}/2\pi$ must be integer, (8.7) leads to the following condition for global wave-particle resonances:

$$\omega = m\Omega. \quad (8.8)$$

This result is, of course, not unexpected. However, note that (8.8) is only a special case that illustrates a more general local resonance condition (8.6), which does not require the wave and particle oscillations to be bounded.

8.2. Metaplectic resonance heating

The metaplectic resonance condition (8.4) also emerges naturally in plasma kinetic theory. Consider particles with a probability distribution function f governed by the Vlasov equation

$$\partial_t f + \{H, f\} = 0, \quad (8.9)$$

where H is the particle Hamiltonian. Suppose $H = H_0 + H_1$, where H_1 is a perturbation small compared with H_0 . Then, within linear theory, one has $f = f_0 + f_1$, where f_0 is the background distribution and f_1 is of order H_1 . These functions satisfy

$$\partial_t f_0 + \{H_0, f_0\} = 0, \quad \partial_t f_1 + \{H_0, f_1\} = -\{H_1, f_0\}. \quad (8.10)$$

Let us assume that H_0 is time-independent and f_0 is an equilibrium distribution $f_0 = f_0(H_0)$. Then, in the canonical variables (τ, H_0) , the Poisson brackets in the equation for f_1 can be expressed as follows:

$$\{H_0, f_1\} = \partial_{\tau} f_1, \quad \{H_1, f_0(H_0)\} = -(\partial_{\tau} H_1) f_0'(H_0). \quad (8.11)$$

⁴⁷Since we treat the particle as a quantum wave, its classical Hamiltonian is $\hbar H_z$ and its classical momentum is $\hbar k$. However, simultaneously rescaling the energy and momentum by \hbar (or any other constant factor) does not affect (8.5), so one can as well interpret H_z in (8.5) as the classical Hamiltonian and z as the classical phase space.

This leads to

$$(\partial_t + \partial_\tau) f_1 = (\partial_\tau H_1) f'_0(H_0). \quad (8.12)$$

Let us assume that the perturbation is an eikonal wave in τ , such that $f_1 = \text{re}\tilde{f}$, $H_1 = \text{re}\tilde{H}$,

$$\partial_t \tilde{f} = -i\omega \tilde{f}, \quad \partial_\tau \tilde{f} = i(\partial_\tau \theta) \tilde{f}, \quad \partial_\tau \tilde{H} = i(\partial_\tau \theta) \tilde{H}. \quad (8.13)$$

Hence,

$$\tilde{f} = -\frac{\partial_\tau \theta}{\omega - \partial_\tau \theta} \tilde{H} f'_0(H_0), \quad (8.14)$$

so particles will be resonant with the wave under the condition (8.4).

As usual, the presence of a linear resonance implies a heating mechanism, which we call metaplectic resonance heating. A quantitative theory of such heating for specific waves is beyond the scope of this paper, but note that the concept of power dissipation is naturally extendable from the physical x -space to metaplectically transformed fields. For example, as shown in [Appendix E](#), there is a remarkably concise formula that connects the energy dissipated by an electromagnetic wave, per volume Γ of the wave's extended phase space, in any linear dielectric medium and plasma, in particular:

$$\frac{d\mathcal{E}}{d\Gamma} = \text{tr}(\sigma_H \mathfrak{W}). \quad (8.15)$$

Here, σ_H is the symbol of the Hermitian part of the conductivity operator, \mathfrak{W} is the average Wigner matrix of the electric field and tr denotes the matrix trace. Traditionally, σ is calculated analytically or semi-analytically using GO ordering. Using MGO, it should be possible to extend those calculations to regions where GO fails, cutoffs in particular.

9. Conclusions

Let us summarise. In this paper, we introduce a formulation of geometrical optics (GO) that is not limited to small λ/L in the physical space. How do we do it? The intrepid readers who hoped to find a short answer here are referred to § 1.3, which is as short as a meaningful answer can be. However, here are some highlights.

As commonly known by now (Tracy *et al.* 2014), the natural language of GO is the Weyl symbol calculus and metaplectic transforms (MTs).⁴⁸ However, neither Weyl symbols nor MTs are inherently tied to the physical space x or the spectral space k for that matter. Thus, there is no fundamental reason, other than tradition, why the envelope dynamics should be calculated in the x -space or the k -space. Furthermore, instead of wavefields, what really matters for practical use are their Wigner functions (or even integrals thereof). As we show, those can be easily recalculated between any two representations using the Airy transform. Then, it makes sense to work in the phase-space variables in which the GO applicability conditions are satisfied best and map the solution to the x -space only occasionally, when needed or convenient.

⁴⁸In fact, one can think of the envelope dynamics itself as a metaplectic image of the dynamics of the complete wavefield, as discussed in § 4.

The most natural variables are the ray time τ and the ray energy h . We explicitly derive the envelope equation in these variables; see §7 for the general case. As usual, the coefficients therein are determined by the Weyl symbol of the dispersion operator, which is the same (within the assumed accuracy) as the corresponding symbol in (x, k) . This means that integrating the envelope equation in MGO is just as easy as integrating the one in GO. Yet, because the derivatives in the MGO equation are taken along cutoff-free directions, the field remains non-singular. As we show, the standard Airy patterns that form in regions where conventional GO fails are successfully reproduced within MGO simply by remapping the field from the τ -space to the x -space.

MGO will likely be useful, for example, for reduced modelling of the O–X conversion in inhomogeneous plasma near the critical density (work in progress). A generalisation to quasioptical wave beams, which can be done mostly like as by Dodin *et al.* (2019) and Yanagihara *et al.* (2019a,b), is left to future work. Overall, MGO can replace GO for any practical purposes, because it better handles cutoffs and is similar otherwise.

Acknowledgements

Editor Alex Schekochihin thanks the referees for their advice in evaluating this article.

Funding

This research was supported by the U.S. Department of Energy through contract No. DE-AC02-09CH11466, by an A*STAR SERC Central Research Fund and by grant NSF PHY-2309135 to the Kavli Institute for Theoretical Physics (KITP). The United States Government retains a non-exclusive, paid-up, irrevocable, world-wide licence to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

Declaration of interests

The authors report no conflict of interest. One of the authors (I.Y.D.) thanks Nicholas Bohlsen for fruitful discussions.

Appendix A. Ray equations

Here, we rederive the ray equations, both for completeness and also to introduce the necessary terminology. In doing so, we generally follow Dodin (2022), but see also Tracy *et al.* (2014). Note that the notation in this appendix differs from that in the main text.

Consider an eikonal wave with a rapid phase $\theta(t, \mathbf{x})$, where t is the physical time and \mathbf{x} is the coordinate in the physical space, which we now allow to be multi-dimensional for generality. The phase determines the wave frequency and wavevector:

$$\bar{\omega}(t, \mathbf{x}) \doteq -\partial_t \theta, \quad \bar{\mathbf{k}}(t, \mathbf{x}) \doteq \nabla \theta. \quad (\text{A1})$$

(We use bars to distinguish these functions from generic coordinates in the spectral space.) Suppose that they satisfy some dispersion relation

$$\Lambda(t, \mathbf{x}, \bar{\omega}, \bar{\mathbf{k}}) = 0, \quad (\text{A2})$$

where Λ is a real function. Let us introduce $w(t, \mathbf{x}, \mathbf{k})$ as the (relevant) solution of

$$\Lambda(t, \mathbf{x}, w(t, \mathbf{x}, \mathbf{k}), \mathbf{k}) = 0 \quad (\text{A3})$$

for a given \mathbf{k} . Differentiating (A3) with respect to t , \mathbf{x} and \mathbf{k} gives

$$\partial_t \Lambda + (\partial_\omega \Lambda) \partial_t w = 0, \quad (\text{A4a})$$

$$\partial_x \Lambda + (\partial_\omega \Lambda) \partial_x w = 0, \quad (\text{A4b})$$

$$\partial_k \Lambda + (\partial_\omega \Lambda) \partial_k w = 0, \quad (\text{A4c})$$

where the derivatives of Λ are evaluated at $(t, \mathbf{x}, w(t, \mathbf{x}, \mathbf{k}), \mathbf{k})$. In particular, (A4c) gives

$$\mathbf{v}_g \doteq \frac{\partial w}{\partial \mathbf{k}} = -\frac{\partial_k \Lambda}{\partial_\omega \Lambda} \quad (\text{A5})$$

for the group velocity \mathbf{v}_g , whose physical meaning is yet to be introduced. The notation $\bar{\mathbf{v}}_g$ used in the following denotes \mathbf{v}_g evaluated on $\mathbf{k} = \bar{\mathbf{k}}(t, \mathbf{x})$.

From (A1), one readily obtains

$$\frac{\partial \bar{k}_i}{\partial t} + \frac{\partial \bar{\omega}}{\partial x^i} = 0, \quad \frac{\partial \bar{k}_j}{\partial x^i} = \frac{\partial \bar{k}_i}{\partial x^j}, \quad (\text{A6})$$

which are known as consistency relations. They lead to

$$\begin{aligned} \left(\frac{\partial}{\partial t} + \bar{\mathbf{v}}_g \cdot \frac{\partial}{\partial \mathbf{x}} \right) \bar{k}_i(t, \mathbf{x}) &= -\frac{\partial w(t, \mathbf{x}, \bar{\mathbf{k}}(t, \mathbf{x}))}{\partial x^i} + \bar{\mathbf{v}}_g \cdot \frac{\partial \bar{k}_i(t, \mathbf{x})}{\partial \mathbf{x}} \\ &= -\left(\frac{\partial w(t, \mathbf{x}, \mathbf{k})}{\partial x^i} \right)_{\mathbf{k}=\bar{\mathbf{k}}(t, \mathbf{x})} - \bar{v}_g^j \frac{\partial \bar{k}_j(t, \mathbf{x})}{\partial x^i} + \bar{v}_g^j \frac{\partial \bar{k}_i(t, \mathbf{x})}{\partial x^j} \\ &= -\left(\frac{\partial w(t, \mathbf{x}, \mathbf{k})}{\partial x^i} \right)_{\mathbf{k}=\bar{\mathbf{k}}(t, \mathbf{x})}, \end{aligned} \quad (\text{A7})$$

and similarly,

$$\begin{aligned} \left(\frac{\partial}{\partial t} + \bar{\mathbf{v}}_g \cdot \frac{\partial}{\partial \mathbf{x}} \right) \bar{\omega}(t, \mathbf{x}) &= \left(\frac{\partial}{\partial t} + \bar{\mathbf{v}}_g \cdot \frac{\partial}{\partial \mathbf{x}} \right) w(t, \mathbf{x}, \bar{\mathbf{k}}(t, \mathbf{x})) \\ &= \left(\frac{\partial w(t, \mathbf{x}, \mathbf{k})}{\partial t} + \bar{v}_g^i \frac{\partial w(t, \mathbf{x}, \mathbf{k})}{\partial x^i} \right)_{\mathbf{k}=\bar{\mathbf{k}}(t, \mathbf{x})} \\ &\quad + \bar{v}_g^i \left(\frac{\partial}{\partial t} + \bar{\mathbf{v}}_g \cdot \frac{\partial}{\partial \mathbf{x}} \right) \bar{k}_i(t, \mathbf{x}) \\ &= \left(\frac{\partial w(t, \mathbf{x}, \mathbf{k})}{\partial t} \right)_{\mathbf{k}=\bar{\mathbf{k}}(t, \mathbf{x})}, \end{aligned} \quad (\text{A8})$$

where we used (A7). Using the convective derivative associated with the group velocity,

$$d/dt \doteq \partial_t + (\bar{\mathbf{v}}_g \cdot \partial_x), \quad (\text{A9})$$

one can rewrite these compactly as

$$\frac{d\bar{k}_i(t, \mathbf{x})}{dt} = - \left(\frac{\partial w(t, \mathbf{x}, \mathbf{k})}{\partial x^i} \right)_{\mathbf{k}=\bar{\mathbf{k}}(t, \mathbf{x})}, \quad \frac{d\bar{\omega}(t, \mathbf{x})}{dt} = \left(\frac{\partial w(t, \mathbf{x}, \mathbf{k})}{\partial t} \right)_{\mathbf{k}=\bar{\mathbf{k}}(t, \mathbf{x})}. \quad (\text{A10})$$

One can represent (A10) as ordinary differential equations for $\bar{\mathbf{k}}(t) \doteq \bar{\mathbf{k}}(t, \bar{\mathbf{x}}(t))$ and $\bar{\omega}(t) \doteq \bar{\omega}(t, \bar{\mathbf{x}}(t))$, where $\bar{\mathbf{x}}(t)$ are the ‘ray trajectories’ governed by

$$\frac{d\bar{x}^i(t)}{dt} = v_{\text{g}}^i(t, \bar{\mathbf{x}}(t), \bar{\mathbf{k}}(t)) \quad (\text{A11})$$

and \mathbf{v}_{g} serves as the ray velocity. Together with (A11), (A10) become Hamilton’s equations:

$$\frac{dx^i}{dt} = \frac{\partial w(t, \mathbf{x}, \mathbf{k})}{\partial k_i}, \quad \frac{dk_i}{dt} = - \frac{\partial w(t, \mathbf{x}, \mathbf{k})}{\partial x^i}, \quad \frac{d\omega}{dt} = \frac{\partial w(t, \mathbf{x}, \mathbf{k})}{\partial t}, \quad (\text{A12})$$

where we have finally dropped the bars for clarity.

The ray equations can also be expressed directly through the dispersion function Λ . To do this, let us introduce a new ray variable σ via

$$\frac{d\sigma}{dt} = - \frac{1}{\partial_{\omega} \Lambda}, \quad (\text{A13})$$

where $\Lambda \equiv \Lambda(t, \mathbf{x}, \omega, \mathbf{k})$. (The minus sign on the right-hand side is a matter of convention.) Then, using (A4), one can write

$$\frac{\partial w}{\partial t} = \frac{\partial \Lambda}{\partial t} \frac{d\sigma}{dt}, \quad \frac{\partial w}{\partial x^i} = \frac{\partial \Lambda}{\partial x^i} \frac{d\sigma}{dt}, \quad \frac{\partial w}{\partial k_i} = \frac{\partial \Lambda}{\partial k_i} \frac{d\sigma}{dt}. \quad (\text{A14})$$

Substituting these into (A12), one finds

$$\frac{dt}{d\sigma} = - \frac{\partial \Lambda}{\partial \omega}, \quad \frac{d\omega}{d\sigma} = + \frac{\partial \Lambda}{\partial t}, \quad (\text{A15a})$$

$$\frac{dx^i}{d\sigma} = + \frac{\partial \Lambda}{\partial k_i}, \quad \frac{dk_i}{d\sigma} = - \frac{\partial \Lambda}{\partial x_i}, \quad (\text{A15b})$$

where we have also included (A13) for completeness. One can view these as Hamilton’s equations with Hamiltonian Λ , canonical pairs $(t, -\omega)$ and (x^i, k_i) , and the ‘ray time’ σ , which is also known as the ray orbit parameter. The ray time coincides with the physical time if the ray Hamiltonian can be represented as

$$\Lambda(t, \mathbf{x}, \omega, \mathbf{k}) = \lambda(t, \mathbf{x}, \mathbf{k}) - \omega. \quad (\text{A16})$$

In the main text, we assume that \mathbf{x} and \mathbf{k} are one-dimensional, $\partial_t \Lambda \equiv 0$, overdots represent derivatives with respect to σ , τ stands for σ expressed as a function of phase-space coordinates, and the physical time t is not used until §8. Then, for $\Lambda = H_z$, (A15b) can be written as (4.11) and (4.10) plays the role of (A2).

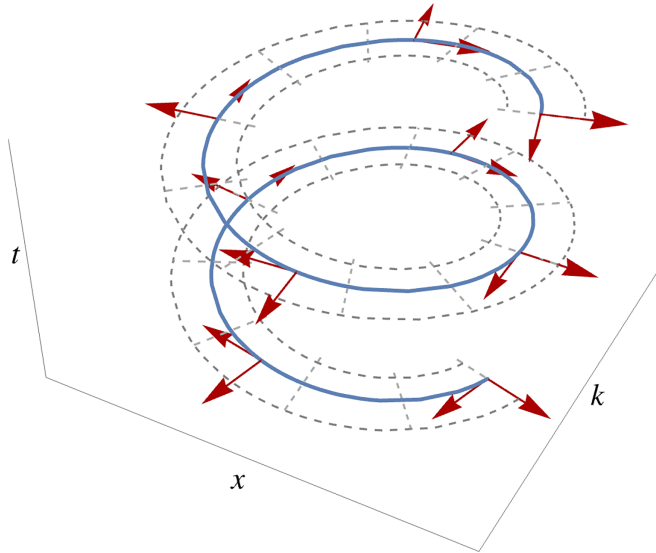


FIGURE 10. Schematic of the local (ϕ, j) coordinate grid for an oscillator: the ray trajectory in the Z -space is shown in blue; the red arrows are the corresponding local basis vectors (5.12); dashed are the local isosurfaces of $\phi(Z)$ and $j(Z)$. Near the ray on the (ϕ, j) -plane, the isosurfaces of j are parallel to the ray and the isosurfaces of ϕ are transverse to the ray. The mapping $(x, k) \mapsto (\phi, j)$ is multi-valued, so ϕ is not restricted to $[-\pi, \pi)$, as a canonical angle normally would, but can range from $-\infty$ to $+\infty$.

Appendix B. Angle–action variables

Here, we elaborate on the idea of adopting MGO variables in the form $(\phi, j)^T \equiv \mathbf{a}$, where ϕ is the angle variable of the OHO, $j \doteq J - J_0$, J is OHO's action and J_0 is OHO's action on the ray. As mentioned in § 5.3, this approach is not optimal. Nevertheless, it is still worth considering in that it helps us introduce certain concepts that are needed in other parts of this paper.

B.1. Operator transformation

Strictly speaking, we are looking for an operator transformation, and the true angle variable cannot be promoted to an operator due to the discreteness of the oscillator's spectrum (Susskind & Glogower 1964).⁴⁹ However, this is not a problem if the angle–action variables are defined only locally and no boundary conditions are imposed on them. By allowing ϕ to be a multi-valued function of Z , one can treat the ϕ -axis as \mathbb{R} even for rays that actually undergo periodic oscillations (figure 10). The action variable can be handled similarly: although the domain of J is only half of the real axis, j is a signed quantity whose large absolute values do not matter, so the j -axis can be treated as \mathbb{R} too. Then, the Weyl symbol calculus can be introduced in the (ϕ, j) -space as usual.

⁴⁹The normalised eigenvectors of \hat{J} with a discrete spectrum satisfy $\langle e_j(J_1) | e_j(J_2) \rangle = \delta_{J_1, J_2}$. This is incompatible with the canonical commutation relation $[\hat{\phi}, \hat{J}] = i$. Indeed, because $\langle e_j(J_1) | [\hat{\phi}, \hat{J}] | e_j(J_2) \rangle = (J_2 - J_1) \langle e_j(J_1) | \hat{\phi} | e_j(J_2) \rangle$, the commutation relation leads to $(J_2 - J_1) \langle e_j(J_1) | \hat{\phi} | e_j(J_2) \rangle = i \delta_{J_1, J_2}$, which cannot be satisfied at $J_1 = J_2$.

The angle–action variables of the OHO are introduced via the parametrisation

$$q = \sqrt{R^2 + 2j} \sin \phi, \quad (\text{B1a})$$

$$p = -R + \sqrt{R^2 + 2j} \cos \phi. \quad (\text{B1b})$$

This leads to the following formulae for (ϕ, j) as functions of (q, p) (for either sign of ε):

$$\phi = \arctan \left(\frac{q}{R + p} \right), \quad (\text{B2a})$$

$$j = \frac{1}{2} (q^2 + (p + R)^2 - R^2). \quad (\text{B2b})$$

We will now seek the corresponding operators such that

$$[\hat{\phi}, \hat{j}] = i. \quad (\text{B3})$$

This means that the corresponding symbols⁵⁰ $\phi(q, p)$ and $j(q, p)$ must satisfy $\{\{\phi, j\}\} = 1$, while (B2) satisfy $\{\phi, j\} = 1$. The Moyal bracket coincides with the Poisson bracket only to the extent that the third-order derivatives of ϕ and j are negligible (§ 2.2), i.e. in the limit of large R . Hence, one can consider (B2) as a large- R asymptotic of the desired transformation, which helps to guess the operators that actually satisfy (B3) with an acceptable accuracy. Specifically, we proceed as follows. In (B2b), we simply replace the symbols q and p with the corresponding operators without making additional changes, i.e. adopt the standard action operator of a harmonic oscillator up to a shift:

$$\hat{j} = \frac{1}{2} (\hat{q}^2 + (\hat{p} + R)^2 - R^2). \quad (\text{B4a})$$

For the angle variable, we consider the large- R asymptotic of (B2a), and then replace q and p with \hat{q} and \hat{p} , correspondingly, in such a way that the resulting operator remains Hermitian (or one can actually take the Weyl transform instead):

$$\hat{\phi} = \frac{\hat{q}}{R} - \frac{\hat{p}\hat{q} + \hat{q}\hat{p}}{2R^2} + \frac{3\hat{p}\hat{q}\hat{p} - \hat{q}^3}{3R^3} + \dots \quad (\text{B4b})$$

Remember that R is the radius of curvature at a fixed point \mathbf{z}_0 , so it is considered independent of ϕ and j in this local approximation and thus commutes with \hat{q} and \hat{p} . By keeping $\mathcal{O}(R^{-m})$ terms in (B4b), one satisfies (B3) up to $\mathcal{O}(R^{-m})$, as can be verified by a direct calculation. Hence, the above-mentioned $\hat{\phi}$ and \hat{j} can be used as base operators. Figure 11 illustrates the construction of the corresponding coordinate mesh, reflecting the evolution of OHO's parameters along the ray.

B.2. M -waves

For the specific transformation considered here, M -waves satisfy the equation of a quantum harmonic oscillator (QHO), $i\partial_\phi \bar{M} = j(q, -i\partial_q) \bar{M}$. The propagator for this

⁵⁰Since the transformation $\mathbf{z} \mapsto \mathbf{y}$ is linear, these can be equally considered as \mathbf{y} -symbols or \mathbf{z} -symbols.

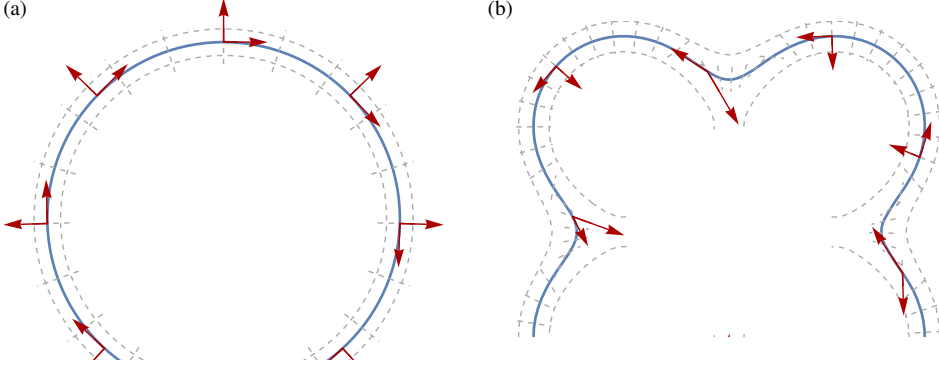


FIGURE 11. Examples of the ray trajectories (blue) in the z -space with the corresponding local basis vectors (5.12) (red arrows). Both figures are produced numerically for sample H_Z . Dashed are the local isosurfaces of $\phi(z)$ and $j(z)$. (a) Constant $g > 0$ and $\varepsilon > 0$; (b) $g < 0$ everywhere, while the sign of ε varies. The directions of the arrows are consistent with those in figure 2.

equation is known exactly (Sakurai & Napolitano 2020, Section 2.6.1). However, the initial conditions for said equation, $\bar{M}(q, 0) = \langle \varepsilon_{\hat{\phi}}(0) | \varepsilon_{\hat{q}}(q) \rangle$, are not, so the exact propagator is not immediately applicable. Because of this, in the following, we use an alternative, asymptotic, approach that follows the general recipe from § 4.3. Let us adopt $\bar{M}(q, \phi) \equiv M^*(\phi, q) = e^{i\Theta(q, \phi)} \mathfrak{M}(q, \phi)$. Then, by (4.30), one has⁵¹

$$\partial_q \Theta(q, \phi) = p(q, \phi) = q \cot \phi - R, \quad (\text{B5a})$$

$$\partial_\phi \Theta(q, \phi) = -j(q, \phi) = \frac{R^2}{2} - \frac{q^2}{2 \sin^2 \phi}, \quad (\text{B5b})$$

where the latter equality is obtained by expressing j from (B1a). This leads to

$$\Theta(q, \phi) = \Theta_0 + \frac{q^2}{2} \cot \phi - Rq + \frac{R^2 \phi}{2}, \quad (\text{B6})$$

where Θ_0 is an integration constant. Then, $\partial_{q\phi} \Theta = -q / \sin^2 \phi$, so (4.40) gives

$$\mathfrak{M}^2(q, \phi) = \frac{|q|}{2\pi \sin^2 \phi}. \quad (\text{B7})$$

The constant phase Θ_0 in (B6) is convenient to choose such that the eigenvectors of \hat{q} and $\hat{\phi}$, $|\varepsilon_{\hat{q}}(q)\rangle$ and $|\varepsilon_{\hat{\phi}}(\phi)\rangle$, coincide at small $q \approx R\phi$ up to normalisation. This corresponds to $M(\phi, q) \approx \sqrt{R} \delta(q - R\phi)$, where the normalisation factor \sqrt{R} is introduced to satisfy (3.21); cf. § 3.5.2. Hence, we adopt

$$M(\phi, q) = \sqrt{\frac{iq}{2\pi \sin^2 \phi}} \exp\left(-\frac{iq^2}{2} \cot \phi + iRq - \frac{iR^2 \phi}{2}\right). \quad (\text{B8})$$

⁵¹Remember that § 4.3 describes a transformation $(x, k) \mapsto (q, p)$, where (q, p) serve the ‘new’ variables. Here, in contrast, we consider a transformation $(q, p) \mapsto (\phi, h)$, so (q, p) serve as the ‘old’ variables. The signs in (B5) have been adjusted accordingly.

Equation (B8) is consistent with (B6) and (B7) and, at small ϕ , gives, as desired,

$$M(\phi, q) \approx \sqrt{\frac{iq}{2\pi\phi^2}} \exp\left(-\frac{i(q-R\phi)^2}{2\phi}\right) \approx \sqrt{R} \delta(q-R\phi), \quad (\text{B9})$$

where the latter equality holds due to the fact that one can replace q with $R\phi$ in the pre-exponential factor and we interpret the phase of the square root as $(\pi/4)\text{sgn}q$. Then, the phase changes by $\pi/2$ across $q=0$, that is, from $-\pi/4$ to $\pi/4$.

The corresponding pseudo-measure (3.29) is given by

$$\begin{aligned} \mu &= \frac{1}{2\pi} \int ds ds' M\left(\phi + \frac{s'}{2}, q + \frac{s}{2}\right) M^*\left(\phi - \frac{s'}{2}, q - \frac{s}{2}\right) e^{-ijs' + ips} \\ &= \frac{1}{(2\pi)^2} \int ds ds' \sqrt{\frac{|(q+s/2)(q-s/2)|}{\sin^2(\phi+s'/2) \sin^2(\phi-s'/2)}} e^{i\varphi}, \\ &= \frac{1}{(2\pi)^2} \int ds ds' \frac{\sqrt{|q^2 - s^2/4|}}{|\sin^2\phi - \sin^2(s'/2)|} e^{i\varphi}, \end{aligned} \quad (\text{B10})$$

where

$$\varphi = -\Theta\left(q + \frac{s}{2}, \phi + \frac{s'}{2}\right) + \Theta\left(q - \frac{s}{2}, \phi - \frac{s'}{2}\right) - js' + ps. \quad (\text{B11})$$

This can be used to recalculate the symbol representation as usual,

$$O_a(\phi, j) = \int dq dp \mu O_y(q, p). \quad (\text{B12})$$

For any smooth symbol O_y , the integral (B12) can be handled like the one in (3.63). That is, one can Taylor-expand Θ in s and s' , keep the linear terms under the exponent, transform the rest and the pre-exponential factor into a polynomial in s and s' , and rewrite it using $s = -i\partial_p e^{ips}$ and $s' = i\partial_j e^{-ijs'}$. This leads to

$$\begin{aligned} O_a(\phi, j) &\approx \frac{1}{(2\pi)^2} \int dq dp ds ds' (1 + \mathcal{O}(s^2, s'^2)) \frac{|q|}{\sin^2\phi} O_y(q, p) e^{i(p-p(q,\phi))s - i(j-j(q,\phi))s'} \\ &= \frac{1}{(2\pi)^2} \int dq dp \frac{|q|}{\sin^2\phi} O_y(q, p) (1 + \mathcal{O}(\partial_p^2, \partial_j^2)) \int ds ds' e^{i(p-p(q,\phi))s - i(j-j(q,\phi))s'} \\ &= \int dq dp \frac{|q|}{\sin^2\phi} O_y(q, p) (1 + \mathcal{O}(\partial_p^2, \partial_j^2)) \delta(j - j(q, \phi)) \delta(p - p(q, \phi)). \end{aligned} \quad (\text{B13})$$

Due to (B5b), one has

$$\frac{|q|}{\sin^2\phi} \delta(j - j(q, \phi)) = \left| \frac{\partial j(q, \phi)}{\partial q} \right| \delta(j - j(q, \phi)) = \delta(q - Q(\phi, j)), \quad (\text{B14})$$

where $Q(\phi, j) \doteq \sqrt{R^2 + 2j} \sin \phi$.⁵² This leads to

$$O_a(\phi, j) = \int dq dp (O_y(q, p) + \mathcal{O}(\partial_y^2 O)) \delta(q - Q(\phi, j)) \delta(p - P(\phi, j)), \quad (\text{B15})$$

where $P(\phi, j) \doteq -R + \sqrt{R^2 + 2j} \cos \phi$. The term $\mathcal{O}(\partial_y^2 O)$ can be estimated as O/R^2 , so it is negligible within the assumed accuracy. Then, unsurprisingly, the leading-order transformation of symbols is simple:

$$O_a(\phi, j) \approx O_y(Q(\phi, j), P(\phi, j)). \quad (\text{B16})$$

However, remember that the approximation (B16) is applicable only to smooth symbols. Also, the usability of (B4b) as an asymptotic series is conditioned upon smallness of κ , which is due to the fact that higher-order derivatives of H_z have been neglected. Also, $\dot{\phi}$ is not globally constant, so (ϕ, j) are not true angle-action variables of the original system even approximately. That is why we prefer the alternative variable transformations discussed in §§ 5.3 and 5.4, which are not associated with such inconveniences.

Appendix C. Auxiliary functions

C.1. Properties of Ai_γ

Here, we discuss some useful properties of the function Ai_γ . By Taylor-expanding the integrand in (5.38) in γ , one obtains

$$\begin{aligned} Ai_\gamma(z) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dz \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{i\gamma t^3}{24} \right)^n e^{izt} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{i\gamma}{24} (-i\partial_z)^3 \right)^n \frac{1}{2\pi} \int_{-\infty}^{\infty} dz e^{izt} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{\gamma}{24} \partial_z^3 \right)^n \delta(z) \\ &= \exp \left(-\frac{\gamma}{24} \partial_z^3 \right) \delta(z). \end{aligned} \quad (\text{C1})$$

In other words,

$$Ai_\gamma(z) = \delta(z) - \frac{\gamma}{24} \delta'''(z) + \dots \quad (\text{C2})$$

Notice that the first two derivatives of the delta function do not contribute to the right-hand side of (C2).

C.2. Properties of \mathfrak{A}

Let us also consider the function \mathfrak{A} defined in (5.58). First of all, a direct calculation shows that

$$\Phi = (H_y(y) - h)s - \frac{is^3}{24} V^3(q) p''(q) - \frac{is^5}{1920} V^5(q) p''''(q) + \dots \quad (\text{C3})$$

⁵²Formally, the equation $j(q, \phi) = j$ has two solutions for q as a function of ϕ at given j , $q = \pm \sqrt{R^2 + 2j} \sin \phi$. The one with the minus sign is ignored here because, as we already mentioned earlier, the MT to the angle-action variables is defined only locally.

Let us rewrite this as

$$\Phi \doteq s \Delta H + \frac{s^3}{24} \varepsilon_2 + \frac{s^5}{1920} \varepsilon_4 + \dots, \quad (\text{C4})$$

where $\Delta H \doteq H_y - h$ and $\varepsilon_n \doteq -V^{n+1} p^{(n)}$. Note that the characteristic scales of q and p are both of order R , and $V \sim v$, so

$$\varepsilon_2 \sim v^3/R \sim \varepsilon, \quad \varepsilon_4 \sim v^5/R^3 \sim g\varepsilon, \quad (\text{C5})$$

where we used $g \sim \Omega^2 \sim v^2/R^2$. The characteristic values of s that contribute to the integral (5.58) are of the order of those at which Φ is stationary, i.e. those that satisfy

$$0 = \partial_s \Phi = \Delta H + \frac{s^2}{8} \varepsilon_2 + \frac{s^4}{384} \varepsilon_4 + \dots \quad (\text{C6})$$

Let us use this to estimate the characteristic scale of \mathfrak{A} with respect to ΔH . Assuming that the nonlinear, in s , terms in (C4) are not too large at values of s that satisfy (C6), one can rewrite (5.58) as follows:⁵³

$$\mathfrak{A}(\Delta H, \dots) = \frac{1}{2\pi} \int ds \left(1 + \frac{s^3}{24} \varepsilon_2 + \frac{s^5}{1920} \varepsilon_4 + \dots \right) e^{is\Delta H} \quad (\text{C7a})$$

$$= \left(1 + \frac{i}{24} \varepsilon_2 \partial_{\Delta H}^3 - \frac{i}{1920} \varepsilon_4 \partial_{\Delta H}^5 + \dots \right) \frac{1}{2\pi} \int ds e^{is\Delta H} \quad (\text{C7b})$$

$$= \left(1 + \underbrace{\frac{i}{24} \varepsilon_2 \partial_{\Delta H}^3}_{\sim \varepsilon_2/\Delta H^3} - \underbrace{\frac{i}{1920} \varepsilon_4 \partial_{\Delta H}^5}_{\sim \varepsilon_4/\Delta H^5} + \dots \right) \delta(\Delta H). \quad (\text{C7c})$$

Then, \mathfrak{A} 's characteristic scale (roughly, the value of ΔH at which \mathfrak{A} has its first zero) is determined by

$$1 + \mathcal{O}(\varepsilon_2/\Delta H^3) + \mathcal{O}(\varepsilon_4/\Delta H^5) + \dots = 0, \quad (\text{C8})$$

and the values of s that matter satisfy $s \sim \Delta H^{-1}$, as seen from the derivation of (C7c). The higher-order terms denoted by the ellipsis can be dealt with just like the third term, so they do not need to be considered separately and are ignored from now on. Then, there are two options.

First, suppose $\varepsilon_2 \ll \varepsilon_4/\Delta H^2$. Then, (C8) yields $\Delta H \sim (g\varepsilon)^{1/5}$ and $s \sim \Delta H^{-1} \sim (g\varepsilon)^{-1/5}$, which is also consistent with (C6). The term $\propto \varepsilon_2$ is negligible when $s^3 \varepsilon_2 \ll 1$, i.e.

$$1 \gg (g\varepsilon)^{-3/5} \varepsilon \sim (\varepsilon^{2/3}/g)^3 \sim R^{4/5}. \quad (\text{C9})$$

Now, suppose $\varepsilon_2 \gg \varepsilon_4/\Delta H^2$. Then, (C8) yields $\Delta H \sim \varepsilon^{1/3}$ and $s \sim \Delta H^{-1} \sim \varepsilon^{-1/3}$, which is also consistent with (C6). The term $\propto \varepsilon_4$ is negligible when $s^5 \varepsilon_4 \ll 1$, i.e.

$$1 \gg \varepsilon^{-5/3} g\varepsilon \sim g/\varepsilon^{2/3} \sim R^{-4/3}. \quad (\text{C10})$$

which is just the condition opposite to (C9). In summary then, \mathfrak{A} is approximately equal to RAF when $R \gg 1$.

⁵³The integral in (C7a) can be recognised as a ‘catastrophe integral’, like those that emerge in theory of caustics. See Lopez & Dodin (2022, (22)) and references therein.

Appendix D. Quantisation condition for closed orbits

For closed orbits, the operators $\hat{\tau}$ and \hat{h} can be introduced much like the angle-action operators in [Appendix B](#), i.e. with τ as a multi-valued function of q and p . Then, one can notice that, actually, multiple parts of the τ -space represent the same parts of the q -space. By doing so, one effectively replaces the commonly used reflecting boundary conditions in the q -space with periodic boundary conditions in the τ -space. This readily yields the familiar quantisation rule for the discrete spectrum of closed orbits as follows.

First of all, $M_{\hat{q} \leftarrow \hat{\tau}}(q, \tau)$ satisfies

$$M_{\hat{q} \leftarrow \hat{\tau}}(0, \tau) \approx M_0 \delta(\tau), \quad (\text{D1})$$

where $M_0 \doteq [\tau'(0)]^{-1/2}$. Accordingly,

$$\psi_{\hat{q}}(0) \approx M_0 \psi_{\hat{\tau}}(0). \quad (\text{D2})$$

However, one can also express $\psi_{\hat{q}}(0)$ through $M_{\hat{q} \leftarrow \hat{\tau}}$ after a full rotation of the ray around the closed orbit, $\tau = T$. The corresponding MT is given by

$$M_{\hat{q} \leftarrow \hat{\tau}}(0, \tau) \approx -M_0 \delta(\tau) \exp(i\Theta_{\circ}), \quad (\text{D3})$$

where the sign change is due to the cumulative Maslov phase $\pm\pi$ that originates from the square root in [\(6.9\)](#) ([§ 4.3](#))⁵⁴ and Θ_{\circ} is the phase change of $M_{\hat{q} \leftarrow \hat{\tau}}$ around the orbit. Then,

$$\psi_{\hat{q}}(0) \approx M_0 \psi_{\hat{\tau}}(T) \exp(-i\pi + i\Theta_{\circ}). \quad (\text{D4})$$

Since $\psi_{\hat{\tau}}$ is conserved on the ray, one has $\psi_{\hat{\tau}}(T) = \psi_{\hat{\tau}}(0)$. Then, by comparing [\(D2\)](#) with [\(D4\)](#), one finds that $-\pi + \Theta_{\circ}$ must be equal to $2\pi n$, where n is integer; i.e. $\Theta_{\circ} = 2\pi n + \pi$. Also, from [\(6.8\)](#), one has $\Theta = \int (p \, dq - h \, d\tau)$ and $h = 0$ on the ray, so

$$\Theta_{\circ} = \oint p \, dq = \frac{1}{2} \oint dy \wedge y = \frac{1}{2} \oint dz \wedge z = \oint k \, dx. \quad (\text{D5})$$

This leads to the well-known Einstein–Brillouin–Keller quantisation rule ([Stone 2005](#)),

$$\oint k \, dx = 2\pi \left(n + \frac{1}{2} \right), \quad (\text{D6})$$

which is also known as the Bohr–Sommerfeld condition in the limit when $1/2$ is negligible compared with n .

Appendix E. Dissipation power

Let us consider a wave with electric field \tilde{E} in a linear medium with a conductivity operator $\hat{\sigma}$. Such a wave induces the current density

$$\tilde{J}(t, \mathbf{x}) = (\hat{\sigma} \tilde{E})(t, \mathbf{x}) \equiv \int dt' \, d\mathbf{x}' \, \underline{\sigma}(t, \mathbf{x}, t', \mathbf{x}') \tilde{E}(t', \mathbf{x}') \quad (\text{E1})$$

⁵⁴Near each zero of said square root, M can be approximated by the asymptotic form [\(B9\)](#) (up to rescaling $d\tau = d\phi/\Omega$). Then, the phase of $M_{\hat{\tau} \leftarrow \hat{q}}$ increases by π over a closed orbit at clockwise direction, so the phase of $M_{\hat{q} \leftarrow \hat{\tau}}$ decreases by π . At counter-clockwise rotation, the phases have the opposite signs, but, in either case, the M -wave changes sign.

and dissipates the energy

$$\mathfrak{E} = \int dt dx \tilde{\mathcal{J}} \cdot \tilde{\mathbf{E}} = \int dt dx \tilde{\mathbf{E}}^\top \hat{\sigma} \tilde{\mathbf{E}} = \int dt dx \tilde{\mathbf{E}}^\dagger \hat{\sigma}_H \tilde{\mathbf{E}}, \quad (\text{E2})$$

where $_H$ denotes the Hermitian part. (We treat $\tilde{\mathbf{E}}$ as a column vector, so $\tilde{\mathbf{E}}^\top$ is a row vector and $\tilde{\mathbf{E}}^\dagger = \tilde{\mathbf{E}}^\top$, because $\tilde{\mathbf{E}}$ is real.) Using (E6) together with (2.53a) from (Dodin 2022), one can rewrite (E2) as follows:

$$\mathfrak{E} = \int dt dx d\omega dk \text{tr}(\sigma_H \star W_{\tilde{\mathbf{E}}}). \quad (\text{E3})$$

Here, tr denotes the matrix trace, \star is the Moyal product and $\sigma \equiv \sigma(t, \mathbf{x}, \omega, \mathbf{k})$ is the Weyl symbol of $\hat{\sigma}$:

$$\sigma(t, \mathbf{x}, \omega, \mathbf{k}) \doteq \int d\tau ds e^{i\omega\tau - ik \cdot s} \underline{\sigma}\left(t + \frac{\tau}{2}, \mathbf{x} + \frac{s}{2}, t - \frac{\tau}{2}, \mathbf{x} - \frac{s}{2}\right), \quad (\text{E4})$$

which satisfies

$$\sigma(t, \mathbf{x}, -\omega, -\mathbf{k}) = \sigma^*(t, \mathbf{x}, \omega, \mathbf{k}), \quad (\text{E5})$$

because $\underline{\sigma}$ is real by definition. Also, $W_{\tilde{\mathbf{E}}} \equiv W_{\tilde{\mathbf{E}}}(t, \mathbf{x}, \omega, \mathbf{k})$ is the Wigner matrix of $\tilde{\mathbf{E}}$:

$$W_{\tilde{\mathbf{E}}} \doteq \frac{1}{(2\pi)^{1+n}} \int d\tau ds e^{-ik \cdot s + i\omega\tau} \tilde{\mathbf{E}}\left(t + \frac{\tau}{2}, \mathbf{x} + \frac{s}{2}\right) \tilde{\mathbf{E}}^\dagger\left(t - \frac{\tau}{2}, \mathbf{x} - \frac{s}{2}\right), \quad (\text{E6})$$

where $n \doteq \dim \mathbf{x}$. By a known theorem,⁵⁵ (E3) can also be written as

$$\mathfrak{E} = \int dt dx d\omega dk \text{tr}(\sigma_H W_{\tilde{\mathbf{E}}}). \quad (\text{E7})$$

Assuming that σ satisfies MGO applicability conditions for the characteristic phase-space scales, i.e. (cf. (5.43))

$$\Delta t \Delta \omega \gg 1, \quad \Delta x \Delta k \gg 1, \quad (\text{E8})$$

one can replace (E7) with

$$\mathfrak{E} = \int \underbrace{dt dx d\omega dk}_{d\Gamma} \text{tr}(\sigma_H \mathfrak{W}), \quad (\text{E9})$$

where \mathfrak{W} is the local average of $W_{\tilde{\mathbf{E}}}$ over a phase-space volume that is large compared with unity yet small compared with the scale of σ_H . As shown in (Dodin, 2022), this averaging makes \mathfrak{W} positive-semidefinite and a local property of a wave (as opposed to the original $W_{\tilde{\mathbf{E}}}$, which can have oscillating tails that can extend over a whole system). Then, one obtains the following phase-space density of \mathfrak{E} :

$$\frac{d\mathfrak{E}}{d\Gamma} = \text{tr}(\sigma_H \mathfrak{W}). \quad (\text{E10})$$

⁵⁵See, for example, Corollary 7 in Appendix B in Ruiz (2017).

To illustrate the connection of (E10) with familiar results, let us consider a medium that is weakly inhomogeneous (if at all) in time and space, so that standard GO applies. Then, the wavefield can be locally represented as $\vec{E} = \text{re}(e^{-i\bar{\omega}t + i\bar{k}\cdot\mathbf{x}} \tilde{\mathbf{e}})$, with constant $\bar{\omega}$, $\bar{\mathbf{k}}$ and $\tilde{\mathbf{e}}$. This corresponds to (Dodin 2022, Appendix A2)

$$\mathfrak{W} = \frac{\tilde{\mathbf{e}}\tilde{\mathbf{e}}^\dagger}{4} \delta(\omega - \bar{\omega}) \delta(\mathbf{k} - \bar{\mathbf{k}}) + \frac{(\tilde{\mathbf{e}}\tilde{\mathbf{e}}^\dagger)^*}{4} \delta(\omega + \bar{\omega}) \delta(\mathbf{k} + \bar{\mathbf{k}}). \quad (\text{E11})$$

Then, the energy dissipated per unit timespace volume $dt \, d\mathbf{x}$ ('dissipation power density') can be expressed as

$$\mathfrak{P} \doteq \int \frac{d\mathfrak{E}}{d\Gamma} d\omega d\mathbf{k} = \frac{1}{2} \tilde{\mathbf{e}}^\dagger \sigma_{\text{H}}(t, \mathbf{x}, \bar{\omega}, \bar{\mathbf{k}}) \tilde{\mathbf{e}}. \quad (\text{E12})$$

To the leading (zeroth) order in $(\Delta t \Delta \omega)^{-1}$, one also has $\sigma_{\text{H}} = (\omega/4\pi) \boldsymbol{\varepsilon}_{\text{A}}$, where $\boldsymbol{\varepsilon}_{\text{A}}$ is the anti-Hermitian part of the symbol of the dielectric operator, or the dispersion tensor. Thus, (E12) leads to the familiar formula (Stix 1992, Section 4.2)

$$\mathfrak{P} = \frac{\bar{\omega}}{8\pi} \tilde{\mathbf{e}}^\dagger \boldsymbol{\varepsilon}_{\text{A}}(t, \mathbf{x}, \bar{\omega}, \bar{\mathbf{k}}) \tilde{\mathbf{e}}. \quad (\text{E13})$$

Appendix F. Selected notation

Symbol	Definition	Properties/explanation
\doteq	definition	
\mapsto	variable transformation	
\Leftrightarrow	Wigner–Weyl isomorphism	§ 2.2
■	placeholder	
$\dot{\blacksquare}$	time derivative on a ray	
$\bar{\blacksquare}$	refers to the inverse transform	
$\hat{\blacksquare}$	operator	
\blacksquare^*	complex conjugate	
\blacksquare^\dagger	Hermitian adjoint	
\blacksquare^\top	transposition	
\blacksquare_0	evaluated at $z = z_0$	
$\blacksquare_{\hat{x}}$	x -representation	§ 2.1
$\blacksquare_{\hat{q}}$	q -representation	§ 2.1
\blacksquare_r	r -symbol	§ 5.1.2
\blacksquare_y	y -symbol	§ 3.4
\blacksquare_z	z -symbol	§ 3.4
$ \blacksquare\rangle$	Dirac ket	vector in \mathbb{H}
$\langle\blacksquare $	Dirac bra	covector in \mathbb{H}
$\blacksquare \cdot \blacksquare$	$\mathbf{a} \cdot \mathbf{b} \doteq \mathbf{a}^\top \mathbf{b}$	$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a} = a_\alpha b^\alpha$
$\blacksquare \wedge \blacksquare$	$\mathbf{a} \wedge \mathbf{b} \doteq \mathbf{a} \cdot \mathbf{J} \mathbf{b}$	$\mathbf{a} \wedge \mathbf{b} = -\mathbf{b} \wedge \mathbf{a}$, $\mathbf{a} \wedge \mathbf{a} = 0$
$\blacksquare \star \blacksquare$	Moyal star	§ 2.2
$[\blacksquare, \blacksquare]$	commutator	
$\{\blacksquare, \blacksquare\}$	Poisson bracket	$\{a, b\} = (\partial_z a) \cdot \mathbf{J}(\partial_z b) = -\{b, a\}$
$\{\{\blacksquare, \blacksquare\}\}$	Moyal bracket	§ 2.2

Symbol	Definition	Properties/explanation
$\hat{1}$	unit operator	
∂	partial derivative	
$\hat{F}, \hat{\Gamma}$	anti-Hermitian part of \hat{D}	§ 4.1
Γ	symbol of \hat{F} , also phase-space volume	§ 4.1 (E9)
Ω	OHO's angular frequency	$\sigma_\varepsilon \sqrt{ g }$, § 5.2
δ	Dirac delta function	
ϵ	MGO parameter	(5.43)
ε	$(\dot{\mathbf{z}} \wedge \dot{\mathbf{z}})_0$	$\varepsilon = -uv = \mathbf{v} \cdot \mathbf{g}\mathbf{v} \equiv \ \mathbf{v}\ _{\mathbf{g}}^2$
θ	eikonal	rapid phase of a GO wave
κ	symplectic curvature	(5.26)
μ	$2\pi W^{(M)}$	(3.29)
σ_{\blacksquare}	sgn \blacksquare	
τ	ray time	§ 5.4.1, Appendix A
$ \mathbf{e}_{\blacksquare}(\blacksquare)\rangle$	normalised eigenvector	$\hat{A} \mathbf{e}_{\hat{A}}(\lambda)\rangle = \lambda \mathbf{e}_{\hat{A}}(\lambda)\rangle$
\mathcal{K}	reference wavevector	$\partial_{\mathbf{z}}\theta$
\mathbb{H}	Hilbert space	
\mathbb{R}	real line	$(-\infty, \infty)$
\mathbf{I}	unit matrix	
\mathbf{J}	canonical symplectic form	(2.22)
\mathbf{S}	symplectic matrix	$\mathbf{S}^\top \mathbf{J} \mathbf{S} = \mathbf{J}$, $\det \mathbf{S} = 1$
$\bar{\mathbf{S}}$	\mathbf{S}^{-1}	$\bar{\mathbf{S}}^\top \mathbf{J} \bar{\mathbf{S}} = \mathbf{J}$, $\det \bar{\mathbf{S}} = 1$
\mathbf{g}	$(\partial_{\mathbf{z}\mathbf{z}}^2 H)_0$	§ 5.1.2
\mathbf{r}	phase-space variables	$(\tau, \mathbf{h})^\top$
$\mathbf{s}_q, \mathbf{s}_p$	columns of $\bar{\mathbf{S}}$	§ 5.1.1
\mathbf{v}	local phase-space velocity	$\mathbf{v} = \dot{\mathbf{z}}_0 = \mathbf{J}(\partial_{\mathbf{z}} H_{\mathbf{z}})_0$
\mathbf{u}	local phase-space acceleration	$\mathbf{u} = \dot{\mathbf{z}}_0 = \mathbf{J}\mathbf{g}\mathbf{v}$
\mathbf{y}	phase-space variables	$(q, p)^\top$
\mathbf{z}	phase-space variables	$(x, k)^\top$
\mathbf{z}_0	selected point in the \mathbf{z} -space	
\mathcal{O}	big O	
Ai_γ	rescaled Airy function	(5.38)
d	differential	
e	Euler's number	$\exp(1)$
i	imaginary unit	$\sqrt{-1}$
$\text{symb } \blacksquare$	Weyl symbol	§ 2.2
$\hat{D}, \hat{\mathbf{D}}$	dispersion operator	
$\hat{H}, \hat{\mathbf{H}}$	Hermitian part of $\hat{D}, \hat{\mathbf{D}}$	§ 4.1
H, \mathbf{H}	symbol of $\hat{H}, \hat{\mathbf{H}}$	
$M(q, x)$	same as $M_{\hat{q} \leftarrow \hat{x}}(q, x)$	§ 3.1

Symbol	Definition	Properties/explanation
$M_{\hat{q} \leftarrow \hat{x}}(q, x)$	MT kernel	$\langle \epsilon_{\hat{q}}(q) \epsilon_{\hat{x}}(x) \rangle$
$\tilde{M}(x, q)$	same as $M_{\hat{x} \leftarrow \hat{q}}(x, q)$	§ 3.1
$M_{\hat{x} \leftarrow \hat{q}}(x, q)$	inverse-MT kernel	$\langle \epsilon_{\hat{x}}(x) \epsilon_{\hat{q}}(q) \rangle$
\hat{O}	generic operator	
O	symbol of \hat{O}	
R	symplectic radius (scale)	§ 5.2 and 5.4.1
$W(\blacksquare)$	Wigner function	§ 2.3
h	ray energy	§ 5.4.1
g	det \mathfrak{g}	§ 5.2
u	$\mathbf{v} \wedge \mathbf{u}/v$	§§ 5.7 and 5.2
v	symplectic speed	$(\epsilon^2/ g)^{1/4}$, § 5.2

Abbreviation	Meaning
GAF	generalised Airy function
GO	geometrical optics
LST	linear symplectic transformation
MGO	metaplectic geometrical optics
MT	metaplectic transform
OHO	osculating harmonic oscillator
QHO	quantum harmonic oscillator
RAF	rescaled Airy function
SE	Schrödinger equation
WKB	Wentzel–Kramers–Brillouin

REFERENCES

ALONSO, M.A. & FORBES, G.W. 1997 Uniform asymptotic expansions for wave propagators via fractional transformations. *J. Opt. Soc. Am. A* **14**, 1279.

ALONSO, M.A. & FORBES, G.W. 1999 New approach to semiclassical analysis in mechanics. *J. Math. Phys.* **40**, 1699.

BALAKIN, A.A., BALAKINA, M.A., PERMITIN, G.V. & SMIRNOV, A.I. 2007 Quasi-optical description of wave beams in smoothly inhomogeneous anisotropic media. *J. Phys. D: Appl. Phys.* **40**, 4285.

BLAIR, M.D. 2022 The Van Vleck formula on Ehrenfest time scales and stationary phase asymptotics for frequency-dependent phases. *Commun. Math. Phys.* **392**, 517.

BLIOKH, K.Y., RODRÍGUEZ-FORTUÑO, F.J., NORI, F. & ZAYATS, A.V. 2015 Spin-orbit interactions of light. *Nat. Photon.* **9**, 796.

CHESTER, C., FRIEDMAN, B. & URSELL, F. 1957 An extension of the method of steepest descents. *Math. Proc. Camb. Phil. Soc.* **53**, 599.

DODIN, I.Y. 2022 Quasilinear theory for inhomogeneous plasma. *J. Plasma Phys.* **88**, 905880407.

DODIN, I.Y. 2024 Quasilinear theory: the lost ponderomotive effects and why they matter. *Rev. Mod. Plasma Phys.* **8**, 35.

DODIN, I.Y. & FISCH, N.J. 2014 Ponderomotive forces on waves in modulated media. *Phys. Rev. Lett.* **112**, 205002

DODIN, I.Y., RUIZ, D.E. & KUBO, S. 2017 Mode conversion in cold low-density plasma with a sheared magnetic field. *Phys. Plasmas* **24**, 122116.

DODIN, I.Y., RUIZ, D.E., YANAGIHARA, K., ZHOU, Y. & KUBO, S. 2019 Quasioptical modeling of wave beams with and without mode conversion. I. Basic Theory. *Phys. Plasmas* **26**, 072110.

DODIN, I.Y. & STARTSEV, E.A. 2021 On applications of quantum computing to plasma simulations. *Phys. Plasmas* **28**, 092101.

- DONNELLY, S.M., LOPEZ, N.A. & DODIN, I.Y. 2021 Steepest-descent algorithm for simulating plasma-wave caustics via metaplectic geometrical optics. *Phys. Rev. E* **104**, 025304.
- ELIASSON, B. & PAPADOPOULOS, K. 2016 HF wave propagation and induced ionospheric turbulence in the magnetic equatorial region. *J. Geophys. Res.: Space Phys.* **121**, 2727.
- FU, Y., DODIN, I.Y. & QIN, H. 2023 Spin hall effect of radiofrequency waves in magnetized plasmas. *Phys. Rev. E* **107**, 055210.
- GALINDO, A. & PASCUAL, P. 1990 *Quantum Mechanics I*. Springer.
- GOLDSTEIN, H., POOLE, C.P. & SAFKO, J.L. 2011 *Classical Mechanics*, 3rd edn. Pearson.
- DE GOSSON, M. 2006 *Symplectic Geometry and Quantum Mechanics*. Birkhäuser Basel.
- HALL-CHEN, V.H., PARRA, F.I. & HILLESHEIM, J.C. 2022 Beam model of Doppler backscattering. *Plasma Phys. Control. Fusion* **64**, 095002.
- HANSEN, F.R., LYNØV, J.P. & MICHELSEN, P. 1985 The O–X–B mode conversion scheme for ECRH of a high-density tokamak plasma. *Plasma Phys. Control. Fusion* **27**, 1077.
- HAYES, W.D. 1973 Group velocity and nonlinear dispersive wave propagation. *Proc. R. Soc. Lond. A Math. Phys. Sci.* **332**, 199.
- HØJLUND MARHOLT, R., SENSTIUS, M.G. & NIELSEN, S.K. 2024 Demonstration of metaplectic geometrical optics for reduced modeling of plasma waves. *Phys. Rev. E* **110**, 025208.
- KAMRAN, N., OLVER, P. & TENENBLAT, K. 2009 Local symplectic invariants for curves. *Commun. Contemp. Math.* **11**, 165.
- KAY, K.G. 1994a Integral expressions for the semiclassical time-dependent propagator. *J. Chem. Phys.* **100**, 4377.
- KAY, K.G. 1994b Numerical study of semiclassical initial value methods for dynamics. *J. Chem. Phys.* **100**, 4432.
- KRAVTSOV, Y.A. & ORLOV, Y.I. 1990 *Geometrical Optics of Inhomogeneous Media*. Springer-Verlag.
- LANGER, R.E. 1937 On the connection formulas and the solutions of the wave equation. *Phys. Rev.* **51**, 669.
- LITTLEJOHN, R. 1986 The semiclassical evolution of wave packets. *Phys. Rep.* **138**, 193.
- LITTLEJOHN, R.G. 1985 Symplectically invariant WKB wave functions. *Phys. Rev. Lett.* **54**, 1742.
- LITTLEJOHN, R.G. & FLYNN, W.G. 1991 Geometric phases in the asymptotic theory of coupled wave equations. *Phys. Rev. A* **44**, 5239.
- LOPEZ, N.A. 2024 Exact boundary-value solution for an electromagnetic wave propagating in a linearly varying index of refraction. *J. Plasma Phys.* **90**, 995900602.
- LOPEZ, N.A. & DODIN, I.Y. 2019 Pseudo-differential representation of the metaplectic transform and its application to fast algorithms. *J. Opt. Soc. Am. A* **36**, 1846.
- LOPEZ, N.A. & DODIN, I.Y. 2020 Restoring geometrical optics near caustics using sequenced metaplectic transforms. *New J. Phys.* **22**, 083078.
- LOPEZ, N.A. & DODIN, I.Y. 2021a Exactly unitary discrete representations of the metaplectic transform for linear-time algorithms. *J. Opt. Soc. Am. A* **38**, 634.
- LOPEZ, N.A. & DODIN, I.Y. 2021b Metaplectic geometrical optics for modeling caustics in uniform and nonuniform media. *J. Opt.* **23**, 025601.
- LOPEZ, N.A. & DODIN, I.Y. 2022 Metaplectic geometrical optics for ray-based modeling of caustics: theory and algorithms. *Phys. Plasmas* **29**, 052111.
- LOPEZ, N.A., HØJLUND, R. & SENSTIUS, M.G. 2024 Regarding the extension of metaplectic geometrical optics to modelling evanescent waves in ray-tracing codes. *Phys. Plasmas* **31**, 083901.
- MADHUSOODANAN, M. & KAY, K.G. 1998 Globally uniform semiclassical wave functions for multidimensional systems. *J. Chem. Phys.* **109**, 2644.
- MARTINEZ, A. 2002 *An Introduction to Semiclassical and Microlocal Analysis*. Springer.
- MASLOV, V.P. & FEDORIUK, M.V. 1981 *Semi-Classical Approximation in Quantum Mechanics*. Reidel.
- MCDONALD, S.W. 1988 Phase-space representations of wave equations with applications to the eikonal approximation for short-wavelength waves. *Phys. Rep.* **158**, 337.
- MELLO, P.A. & MOSHINSKY, M. 1975 Nonlinear canonical transformations and their representations in quantum mechanics. *J. Math. Phys.* **16**, 2017.
- MILLER, W.H. 1974 Classical-limit quantum mechanics and the theory of molecular collisions. In *Advances in Chemical Physics* (ed. I. Prigogine & S.A. Rice). Wiley.

- MJØLHUS, E. 1990 On linear conversion in a magnetized plasma. *Radio Sci.* **25**, 1321.
- MOSTOWSKI, J. & PIETRASZEWICZ, J. 2021 Wigner function for harmonic oscillator and the classical limit. arXiv: [2104.06638](https://arxiv.org/abs/2104.06638).
- NICACIO, F. 2021 Williamson theorem in classical, quantum, and statistical physics. *Am. J. Phys.* **89**, 1139.
- NOVIKAU, I., DODIN, I.Y. & STARTSEV, E.A. 2023 Simulation of linear non-Hermitian boundary-value problems with quantum singular-value transformation. *Phys. Rev. Appl.* **19**, 054012.
- NOVIKAU, I., STARTSEV, E.A. & DODIN, I.Y. 2022 Quantum signal processing for simulating cold plasma waves. *Phys. Rev. A* **105**, 062444.
- PEREVERZEV, G.V. 1998 Beam tracing in inhomogeneous anisotropic plasmas. *Phys. Plasmas* **5**, 3529.
- POLI, E. *et al.* 2018 TORBEAM. 2.0, a paraxial beam tracing code for electron-cyclotron beams in fusion plasmas for extended physics applications. *Comput. Phys. Commun.* **225**, 36.
- PREINHAELTER, J. & KOPECKÝ, V. 1973 Penetration of high-frequency waves into a weakly inhomogeneous magnetized plasma at oblique incidence and their transformation to Bernstein modes. *J. Plasma Phys.* **10**, 1.
- RUIZ, D.E. 2017 Geometric theory of waves and its applications to plasma physics. PhD thesis, Princeton University.
- RUIZ, D.E. & DODIN, I.Y. 2015 First-principles variational formulation of polarization effects in geometrical optics. *Phys. Rev. A* **92**, 043805.
- RUIZ, D.E. & DODIN, I.Y. 2017*a* Extending geometrical optics: a Lagrangian theory for vector waves. *Phys. Plasmas* **24**, 055704.
- RUIZ, D.E. & DODIN, I.Y. 2017*b* Ponderomotive dynamics of waves in quasiperiodically modulated media. *Phys. Rev. A* **95**, 032114.
- SAKURAI, J.J. & NAPOLITANO, J. 2020 *Modern Quantum Mechanics*, 3rd edn. Cambridge University Press.
- SILVA, D.J. 2007 An accuracy improvement in Egorov's theorem. *Publ. Mat.* **51**, 77.
- STIX, T.H. 1992 *Waves in Plasmas*, 2nd edn. AIP.
- STONE, A.D. 2005 Einstein's unknown insight and the problem of quantizing chaos. *Phys. Today* **58**, 37.
- SUSSKIND, L. & GLOGOWER, J. 1964 Quantum mechanical phase and time operator. *Physics* **1**, 49.
- 'T HOOFT, G. 2024 The hidden ontological variable in quantum harmonic oscillators. arXiv: [2407.18153](https://arxiv.org/abs/2407.18153).
- THOMSON, C.J. & CHAPMAN, C.H. 1985 An introduction to Maslov's asymptotic method. *Geophys. J. Astr. Soc.* **83**, 143.
- TRACY, E.R., BRIZARD, A.J., RICHARDSON, A.S. & KAUFMAN, A.N. 2014 *Ray Tracing and Beyond: Phase Space Methods in Plasma Wave Theory*. Cambridge University Press.
- TRACY, E.R., KAUFMAN, A.N. & BRIZARD, A.J. 2003 Ray-based methods in multidimensional linear wave conversion. *Phys. Plasmas* **10**, 2147.
- VALLÉE, O., SOARES, M. & DE IZARRA, C. 1997 An integral representation for the product of Airy functions. *Z. Angew. Math. Phys.* **48**, 156.
- WEINBUB, J. & FERRY, D.K. 2018 Recent advances in Wigner function approaches. *Appl. Phys. Rev.* **5**.
- WIDDER, D.V. 1979 The Airy transform. *Am. Math. Mon.* **86**, 271.
- YANAGIHARA, K., DODIN, I.Y. & KUBO, S. 2019*a* Quasioptical modeling of wave beams with and without mode conversion. II. Numerical simulations of single-mode beams. *Phys. Plasmas* **26**, 072111.
- YANAGIHARA, K., DODIN, I.Y. & KUBO, S. 2019*b* Quasioptical modeling of wave beams with and without mode conversion. III. Numerical simulations of mode-converting beams. *Phys. Plasmas* **26**, 072112.
- YANAGIHARA, K., KUBO, S. & DODIN, I.Y. & the LHD Experiment Group 2021 Quasioptical propagation and absorption of electron cyclotron waves: simulations and experiment. *Nucl. Fusion* **61**, 106012.
- ZHU, H. & DODIN, I.Y. 2021 Wave-kinetic approach to zonal-flow dynamics: recent advances. *Phys. Plasmas* **28**, 032303.
- ZIOLKOWSKI, R.W. & DESCHAMPS, G.A. 1984 Asymptotic evaluation of high-frequency fields near a caustic: an introduction to Maslov's method. *Radio Sci.* **19**, 1001.
- ZOR, D. & KAY, K.G. 1996 Globally uniform semiclassical expressions for time-independent wave functions. *Phys. Rev. Lett.* **76**, 1990.